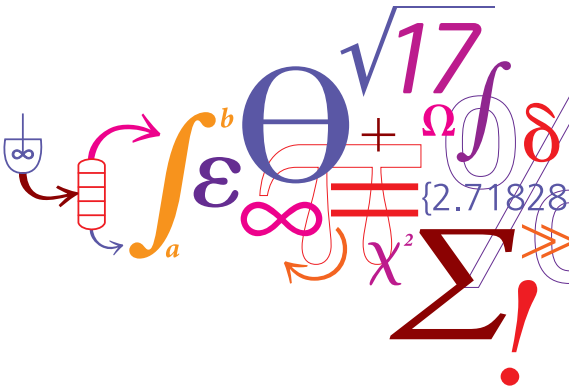
Modelling and operation of reactors for enzymatic biodiesel production

Price, Jason Anthony

Publication date:
2014

Document Version
Publisher's PDF, also known as Version of record

Modelling and Operation of Reactors for Enzymatic Biodiesel Production



Jason Price
Ph.D. Thesis
September 2014

DTU Chemical Engineering
Department of Chemical and Biochemical Engineering

Technical University of Denmark

Modelling and Operation of Reactors for Enzymatic Biodiesel Production

2014

Ph.D Thesis

DTU Chemical Engineering
Department of Chemical and Biochemical Engineering

Modelling and Operation of Reactors for Enzymatic Biodiesel Production

September 2014

Ph.D Thesis

Supervisors:
John M. Woodley
Jakob K. Huusom
Mathias Nordblad

Address: CAPEC-PROCESS
Computer Aided Process Engineering/
Process Engineering and Technology center

Department of Chemical and Biochemical Engineering
Technical University of Denmark
Building 229
DK-2800 Kgs. Lyngby
Denmark
Phone: +45 4525 2800
Fax: +45 4593 2906
Web: www.capec-process.kt.dtu.dk

Print: J&R Frydenberg A/S
København
January 2015

ISBN: 978-87-93054-58-5

Preface

This thesis is submitted as partial fulfilment of the requirements for the Doctor of Philosophy (PhD) degree at the Technical University of Denmark (DTU). The PhD-project has been carried out at the CAPEC-Process Centre at the Department of Chemical and Biochemical Engineering from September 2011 until September 2014 under the supervision of Professor John M. Woodley, Assistant Professor Jakob K. Huusom and Senior Researcher Mathias Nordblad. The project was partly funded by the Danish National Advanced Technology foundation whose support is gratefully acknowledged.

Firstly, I would like to thank my supervisors for giving me the opportunity to grow as an individual and to work on a very interesting and challenging project. Their valuable support, feedback, and easy going attitude throughout my project was greatly appreciated. I would also like to thank my collaborators Per Munk Nielsen and Anders Rancke-Madsen from Novozymes whose project discussions were always stimulating and helped me to focus on the industrial relevance of my project. Thanks to our secretaries Eva and Gitte for all your help with administrative matters. Many thanks, to all my co-workers/friends for providing this wonderful atmosphere in which it was easy to work and for all the shared social events inside and outside DTU. Special mention has to be made to my office mates Rolf, Krešimir and Asbjørn who provided good stimulating conversations, intellectual and otherwise.

I must acknowledge Professor John Villsadsen who encouraged me to peruse my studies in at DTU, and whose hospitality will never be forgotten. A special thanks to my parents for their support and guidance throughout my studies. Last but for sure not least, thanks to my wonderful wife who gave me her unconditional and constant support, for making me laugh, not have to cook and for being my solace through it all.

Kgs. Lyngby, September 2014

Jason Price

Abstract

In developing sustainable industrial processes, biochemical engineering, as a part of a broader field of chemical engineering is becoming an increasingly important as a tool in the chemical engineers toolbox. Its application is driven by consumer demand for new products and by industry wishing to increase profits while reducing operating cost, as well as meeting government and regulatory pressures for processes to be environmentally friendly and sustainable. Current applications of biocatalysts, more specifically, enzymes for large scale bulk production of chemicals have been successfully applied to the production of high fructose corn syrup, upgrading of fats and oils and biodiesel production to name a few. Despite these examples of industrial enzymatic applications, it is still not “clear cut” how to implement biocatalyst in industry and how best to optimize the processes. This is because the processing strategy is usually different to most traditional catalytic processes. In nature, enzymes operate at much lower substrate and product concentrations compared to most industrial chemical processes. What this means is that the natural conditions for biocatalysts are normally much different from conventional process-relevant conditions. Also, the optimal process conditions can vary greatly from one biocatalyst to the next. Hence, to maximize product yields and reactor productivity then the type of reactor operation and downstream processing need to be able to address the aforementioned issues. One way to achieve this is through process modelling to help focus the experimental work needed for process understanding and to support further process development and optimization of the process.

To address how the reactors should be operated; a strategy using mechanistic modelling by combining the biological aspects of the enzyme with reaction/reactor engineering is performed. This strategy is applied to a case study of biodiesel production catalysed by a liquid enzyme formulation. The use of enzymes for biodiesel production is still in its infancy with non-optimized process designs. Furthermore is it unclear how the process should be operated to ensure optimal economics given the relatively high cost of the enzyme and the low value of the products.

In this thesis, the developed mechanistic kinetic model for the enzyme catalysed biodiesel production is used to guide the experimental work. Using the developed model, the prediction and validation of an optimal methanol feeding for Fed-batch operation is achieved along with strategy with moving from fed-batch operation to continuous operation using a liquid lipase. Also addressed is the mismatch between the process and model data given that it is not possible to capture all the underlying phenomena of the process. State estimation theory is used where experimental data is coupled with the developed kinetic model to aid in correcting for the process-model mismatch.

It is shown in this study that the use of conventional chemical engineering principles work aptly well for bio-catalytic processes. While the enzymatic biodiesel case is a “special” in terms of the enzyme being able to operate at much higher substrate and product concentrations the workflow is still valid for other bio-catalytic processes. The modelling of the system initially can be time intensive. However, it pays great dividends given that it gives one the ability to quickly evaluate the applicability of a particular biocatalyst in an industrial process; as well as the ability to quickly evaluate the various reactor configurations to reach the desired process metrics.

Dansk Resumé

I udviklingen af mere bæredygtige industrielle processer spiller biokemiske ingeniørværktøjer, som en del af det generelle område indenfor kemiteknik, en stigende rolle. Denne udviklingen er drævet af et stigende markedet for nye produkter, industriens ønske om øget profit ved at sænke produktionsomkostningerne samt politiske og lovgivningsmæssige krav om mere miljøvenlige og bæredygtige produktionsprocesser. Nuværende applikationer af biokatalyse, specielt for storskala produktioner, er enzymatiske processer for f.eks. produktion af kemikalier fra majs sirup med højt fruktose indhold, opgraderingen af fedt og olier og produktion af biodiesel. På tros af disse eksempler på industrielle enzymatiske processer så er processen med at implementere biokatalyse i industrien ikke moden og der er ikke systematiske metoder til at optimere processerne. Dette skyldes bl.a. at strategier til udvikling af biokatalytiske processer er forskellige for dem som benyttes til udvikling af traditionelle katalytiske processer. Enzymer er aktive ved meget lavere substrat og produkt koncentrationer sammenlignet med typiske industrielle processer. Det betyder at procesbetingelserne er væsentligt anderledes for biokatalytiske processer. Yderligere kan de optimale procesbetingelser ændre sig væsentligt fra en biokatalytisk proces til en anden. Dette forhold er væsentligt at tage i betragtning for at maksimere reaktor udbyttet, ved design af typen af reaktor og oprensningsprocesserne. En metode til at opnå dette maksimale udbytte på er gennem modellering som værktøj til at lede det eksperimentelle arbejde som behøves til at udvikle den nødvendige procesforståelse og optimering.

For at kunne forudsige hvordan en biokatalytisk reaktor skal opereres er mekanistiske modeller i kombination med biokemisk reaktionsteknik benyttet. Denne fremgangmåde er anvendt i et studie af produktion af biodiesel men en flydende formulering af enzymet som biokatalysator. Kommerciel udnyttelse af enzymer til produktion af biodiesel er i sin udviklingsfase og der eksistere ikke et fuldt optimeret procesdesign endnu. Ydereligere er det usikkert hvordan denne proces opereres optimalt i økonomisk henseende grundet de høje omkostninger til enzymet i kombination med en lav pris på produktet.

I denne afhandling er der udviklet en mekanistisk model for produktionen af enzymatisk biodiesel som er benyttet til at planlægge det eksperimentelle arbejde. Med udgangspunkt i modellen er de optimale driftsbetingelser for tilsætningen af metanol blevet bestemt og valideret. Der er yderligere udviklet en strategi for at udvikle en kontinuert proces på basis af et design med fed-batch reaktorer. Der er angivet hvorledes de uundgåelige forskelle der må være mellem den mekaniskske model og den reale reaktor opførsel kan adresseres. Ved hjælp af en tilstandsestimations algoritme kan operations data kombineres med simuleringsmodellen og dermed korrigere for fejl i forudsigelserne.

Det er vist gennem dette studie at gennerelle principper for udvikling af kemiske processer også virker for udviklingen af biokatalytiske processer. Det noteres dog at biodiesel produktionen er et specielt tilfælde inden for klasses af biokatalytiske processer, da det i dette tilfælde er muligt at opererer under relativt høje substrat og produkt koncentrationer. Det menes dog konklusionerne på basis af dette eksempel er af generel karakter. Selvom modellering af en proces kan være et langsomligt stykke arbejde, så betaler det sig idet der er væsentligt hurtigere at evaluere anvendeligheden af en given biokatalysator i en industriel proces. Dette sker igennem en hurtigere evaluering af forskellige konfigurationer af reaktoren i forhold til at opnå de givende processpecifikationer.

List of Papers and Declaration of Contribution

Paper I - Application of Uncertainty and Sensitivity Analysis to a Kinetic Model for Enzymatic Biodiesel Production

Price, J. A., Nordblad, M., Woodley, J., & Huusom, J. K.

Proceedings of 12th IFAC Symposium on Computer Applications in Biotechnology, 161-168(2013)

The project idea was from me and I did all the practical work. The first version of the manuscript was written by me and finalised in collaboration with Mathias Nordblad (MAN), Jakob Huusom (JKH) and John Woodley (JW)

Paper II - . Mechanistic Modelling of Biodiesel Production using a Liquid Lipase.

Price, J., Hofmann, B., Silva, V. T. L., Nordblad, M., Woodley, J. M., & Huusom, J. K.

Biotechnology Progress (2014), doi:10.1002/btpr.1985

The project idea was from MAN and me. Myself, MAN and Bjorn Hofmann(BH) worked on the mechanism. I, BH and Vanessa Silva performed the experimental work. The first version of the manuscript was written by me and finalised in collaboration with MAN, JKH and JW.

Paper III - Fed-Batch Feeding Strategies for Enzymatic Biodiesel Production.

Price, J. A., Nordblad, M., Woodley, J., & Huusom, J. K.

Proceedings of 19th World Congress of the International Federation of Automatic Control, 6204-6209(2013)

The project idea was from me and I did all the practical work. The first version of the manuscript was written by me and finalised in collaboration with MAN, JKH and JW.

Paper IV - From Fed-Batch to Continuous Enzymatic Biodiesel Production

Price, J. A., Nordblad, M., Woodley, J., & Huusom, J. K. (2014).

Manuscript in preparation

The project idea was from Per Munk N Neilsen (Novozymes) and me. I did all the practical work. The first version of the manuscript was written by me and finalised in collaboration with MAN, JKH and JW.

Paper V - Real-Time Model Based Process Monitoring of Enzymatic Biodiesel Production.

Price, J. A., Nordblad, M., Woodley, J., & Huusom, J. K.

Biotechnology Progress (Submitted)

The project idea was from JKH and I did all the practical work. The first version of the manuscript was written by me and finalised in collaboration with MAN, JKH and JW.

Table of Contents

Chapter 1: Introduction

The general thesis structure is outlined along with what makes bio-catalytic processes interesting and where this thesis fits into addressing the modelling and reactor operation for an industrial enzymatic process (Enzymatic biodiesel production).

1.1. Industrial Bio-catalysis

We are in a significant period in time, where in the development of industrial processes we strive to encompass the entire life cycle of a process, to ensure that it is sustainable and environmentally friendly. It goes without saying that these types of processes also need to be economically competitive compared to the conventional process. One of the philosophies in chemical engineering design that has made a significant impact in the way how we design processes is the 12 principles of Green Chemistry[1]. The main outcome from following these principles in the design of products and processes is that it should eliminate hazardous chemicals, minimize waste generation and energy consumption. Bioengineering (from genetic engineering of the biocatalyst to the use of the biocatalyst in a process) is one area of research seen as a promising technique for achieving these green chemistry goals. Biocatalysts are conventionally divided into enzymes and whole cells[2]. In reality, this reflects a spectrum of entities from growing cells to purified enzymes; which themselves may be used in free form or immobilised to facilitate reuse of the biocatalyst[3]. It has been estimated that around 150 bio-catalytic processes based on enzymes have been implemented industrially[4]. Common to all these processes is that a biocatalyst is advantageous for processes where:

- Conventional catalysis pose challenges
- Selectivity needs to be enhanced
- Milder operating conditions are advantageous

Interestingly enough it has been estimated that the industrial enzyme market is valued at around €2.6bn, around 10% of the total catalyst market[4]. However the main drawback related to the use of enzymes are the usually low product yields and poor enzyme stability at relevant process conditions[5,6]. Extensive work in the genetic modification of the biocatalyst and the development of suitable bio-catalytic carriers has helped to raise bio-catalytic productivity (g product/g Enzyme) and increased the enzyme stability[7–9]. Likewise, modifying the process by using engineering design techniques such as in-situ product removal and substrate feeding strategies have also proven beneficial in making bio-catalytic processes a reality and pushing the operating boundaries of the biocatalysts[10–13]. It is within this area of engineering design where this research is based.

1.2. The role of modelling for implementation of bio-catalysis

The focus of this work is on the upstream processing where the enzyme catalysed reaction is taking place, more specifically the reactor operation. Most bio-catalytic reactions that exhibit substrate inhibition or deactivation by high concentrations of the substrate are operated as fed-batch so as to mitigate these effects. However for high volume production a continuous system is desirable. The substrate feeding strategy to maximise the plant productivity in the Fed-batch case and how to operate the continuous process is not a straightforward process.

What we hypothesize is that by using a mechanistic modelling approach that the knowledge gained will give insight into how the reactor configuration should be designed and operated for a bio-catalytic process. Much work has been placed on mechanistic modelling and simulation of bio-catalytic systems[14–19]. What one will notice is that for all the bio-catalytic process mentioned, the kinetic model is combined with the mass balance for the system. However, the model of the system is never extended to a different reactor system. For most cases this is fine given the model is developed to fit a specific case or purpose. However, there are a number of issues needed to be resolved in the modelling and operation of enzymatic processes. In relation to the modelling of enzymatic processes, given that the reactions are complex (parallel and sequential reactions taking place at the same time) and that the system is usually multiphasic; the usual challenge is to:

- Decide on the model complexity
- Identify the parameters for the model of the system.
- Deal with identifiability issues found and still use the developed process model for predictive purposes.

There is a significant body of scientific literature on the modelling and identifiability of bio-catalytic systems[14,20–25], but none of these use the difference in the mass balance for the different types of reactors (example batch vs CSTR) to aid in the identifiability of the system. To investigate the modelling and operation of reactors for bio-catalytic reactions, we chose an industrially relevant process that satisfies the three criteria for where a biocatalyst is advantageous, Enzymatic Biodiesel Production.

1.3. Enzymatic Biodiesel Production as a Case Study

Interest in the production of renewable fuels coupled with environmental concerns, mainly due to global warming, has led to increased research into the production of biofuels, such as biodiesel[26–28]. Oils and fats, which are too viscous to be used directly in engines, are converted into their corresponding methyl or ethyl esters by a process called transesterification[29]. The chemical catalyst route reacts the oil/fat (mainly composed of acyl glycerides and free fatty acids) with alcohol (mainly methanol), in the presence of a strong catalyst (e.g. sodium hydroxide). This results in the formation of biodiesel (fatty acid alkyl esters) together with glycerol as a by-product[30].

Many of the drawbacks associated with chemical catalysis, such as, not being able to easily treat feeds with a high free fatty acid content, recovery of high purity glycerol and high excess methanol input can be overcome by using lipases (triacylglycerol acylhydrolase, EC 3.1. 1.3) as a biocatalyst for transesterification [31,32]. It is well documented that enzymatic processing of oils and fats for biodiesel is technically feasible [33–35]. However, with very few exceptions, a biocatalyst is not the main "go to" catalyst for commercial-scale biodiesel production. This is mainly due to the maturity of the technology, non-optimized process design, and a lack of available cost effective enzymes. Given the relatively high cost of the bio-catalyst, early biodiesel processes employed immobilised enzymes for easy recovery and reuse of the enzyme. More recently it has been reported that the biocatalyst cost can be reduced by using a liquid lipase formulation[36]. However, there has not been much work on how this process should be operated when using a liquid lipase formulation.

Likewise, in terms of process design and synthesis of biodiesel, what is particularly challenging when using an enzymatic catalyst in that to thermodynamically shift the reaction to favour higher biodiesel yields excess methanol is used[30]. However, this methanol can inhibit and inactivate the enzymes. This means that it is essential to identify the reaction conditions that allow for the optimal catalytic rate and enzymatic stability. The choice between different types and combinations of reactors and separation units is greatly affected by the range of operating conditions at which the productivity goals can be met. Thus making detailed knowledge of these parameters essential to the overall process design[37]. Furthermore, as a low priced bulk chemical, given the narrow operating margins, it

is unclear how this process needs to be monitored and operated to ensure optimal economics. Devising a strategy for operating an enzymatic biodiesel process is therefore essential. In devising an operating strategy we see mechanistic model based design as one way to gain better understanding of the process.

1.4. Project Objectives

Many mechanistic models for enzymatic biodiesel production have been proposed but they usually fall short in terms of using the said model for process design purposes; such as, predicting of an optimal methanol feeding profile or how a continuous process should be operated[18,19,38–42].The aim of this research is then the strategy used in model based process design to aid in the operation and development of enzymatic biodiesel production. The mechanistic model of the process is used to evaluate how to operate an enzymatic biodiesel process using a liquid lipase for fed-batch and continuous operation. In terms of reactor operation, while mixing and temperature are important, these are usually fixed and the focus is placed on substrate feeding given it is reported that the feedstock represents over 85 % of the biodiesel production cost[36]. Hence efficient use of the substrate is essential. It is then envisaged that the tools and methods used could be applied to other bioprocesses given the general workflow used in the thesis.

1.5. Thesis Structure and Content

The structure, organization and content of this PhD thesis is visualized in the flow chart presented in Figure 1-1. The various chapters can be grouped together according to various themes. Chapter 2 and 3 give the general background on enzymatic biodiesel production. Chapter 4 to 6 give details on the methodology and tools used along with the model development, while chapters 7 to 9 show the application of the developed mechanistic model for various process design cases. Chapter 10 ties the work together and draws more general conclusions about the material presented in Chapters 5 to 9.

The selected published/submitted journal and reviewed conference papers are also illustrated in the flow chart. This thesis therefore complements the already published papers by providing a condensed and coherent presentation of the overall project and related results.

Figure 1-1 Thesis structure and content

1.6. References

1. Anastas, Paul T and Warner JC, Anastas PT, Warner JC. *Green Chemistry: Theory and Practice*. Oxford University Press; 2000.

2. Gardossi L, Poulsen PB, Ballesteros A, Hult K, Švedas VK, Vasić-Rački Đ, Carrea G, Magnusson A, Schmid A, Wohlgemuth R, Halling PJ. Guidelines for reporting of biocatalytic reactions. *Trends Biotechnol*. 2010;28:171–180.

3. Halling PJ. Thermodynamic predictions for biocatalysis in nonconventional media: Theory, tests, and recommendations for experimental design and analysis. *Enzyme Microb Technol*. 1994;16:178–206.

4. The enzyme makers. *Chem Ind*. 2012;76(11):28-31.

5. Bassegoda A, Cesarini S, Diaz P. Lipase improvement: goals and strategies. *Comput Struct Biotechnol J*. 2012;2(3):1-8.

6. Tufvesson P, Lima-Ramos J, Nordblad M, Woodley JM. Guidelines and Cost Analysis for Catalyst Production in Biocatalytic Processes. *Org Process Res Dev*. 2011;15:266-274.

7. Pollard DJ, Woodley JM. Biocatalysis for pharmaceutical intermediates: the future is now. *Trends Biotechnol*. 2007;25:66–73.

8. Nordblad M, Adlercreutz P. Immobilisation procedure and reaction conditions for optimal performance of Candida antarctica lipase B in transesterification and hydrolysis. *Biocatal Biotransformation*. 2013;31:237-245.

9. Brask J. Immobilized Enzymes in Organic Synthesis. In: *Power of Functional Resins in Organic Synthesis*. Weinheim, Germany; 2008:365 - 380.

10. Woodley JM, Bisschops M, Straathof AJJ, Ottens M. Future directions forin-situ product removal (ISPR). *J Chem Technol Biotechnol*. 2008;83(2):121-123.

11. Freeman A, Woodley JM, Lilly MD. In Situ Product Removal as a Tool for Bioprocessing. *Bio/Technology*. 1993;11(9):1007-1012.

12. Lynch RM, Woodley JM, Lilly MD. Process design for the oxidation of fluorobenzene to fluorocatechol by Pseudomonas putida. *J Biotechnol*. 1997;58(3):167-175.

13. Albaek MO, Gernaey K V, Hansen MS, Stocks SM. Modeling enzyme production with Aspergillus oryzae in pilot scale vessels with different agitation, aeration, and agitator types. *Biotechnol Bioeng*. 2011;108(8):1828-40.

14. Gernaey K V, Lantz AE, Tufvesson P, Woodley JM, Sin G. Application of mechanistic models to fermentation and biocatalysis for next-generation processes. *Trends Biotechnol*. 2010;28(7):346-54.

15. Vasić-Rački D, Findrik Z, Vrsalović Presečki A. Modelling as a tool of enzyme reaction engineering for enzyme reactor development. *Appl Microbiol Biotechnol*. 2011;91(4):845-56.

16. Vasic-Racki D, Kragl U, Liese A. Benefits of Enzyme Kinetics Modelling *. *Chem Biochem Eng Q*. 2003;17(1):7-18.

17. Van Hecke W, Bhagwat A, Ludwig R, Dewulf J, Haltrich D, Van Langenhove H. Kinetic modeling of a bi-enzymatic system for efficient conversion of lactose to lactobionic acid. *Biotechnol Bioeng*. 2009;102(5):1475-82.

18. Al-Zuhair S. Production of Biodiesel by Lipase-Catalyzed Transesterification of Vegetable Oils: A Kinetics Study. *Biotechnol Prog*. 2005;21:1442–1448.

19. Fedosov SN, Brask J, Pedersen AK, Nordblad M, Woodley JM, Xu X. Kinetic model of biodiesel production using immobilized lipase Candida antarctica lipase B. *J Mol Catal B Enzym*. 2013;85-86:156-168.

20. Rodriguez-Fernandez M, Banga JR, Doyle FJ. Novel global sensitivity analysis methodology accounting for the crucial role of the distribution of input parameters: application to systems biology models. *Int J Robust Nonlinear Control*. 2012;22:1082–1102.

21. Sin G, Gernaey K V, Neumann MB, van Loosdrecht MCM, Gujer W. Global sensitivity analysis in wastewater treatment plant model applications: Prioritizing sources of uncertainty. *Water Res*. 2011;45:639-651.

22. Yue H, Halling P, Yu H. Model Development and Optimal Experimental Design of A Kinetically Controlled Synthesis System. In: *Proceedings of 12th IFAC Symposium on Computer Applications in Biotechnology*.; 2013:332-337.

23. Yue H, Brown M, He F, Jia JF, Kell DB. Sensitivity Analysis and Robust Experimental Design of a Signal Transduction Pathway System. *Int J Chem Kinet*. 2008;40:730-741.

24. Brun R, Kuhni M, Siegrist H, Gujer W, Reichert P. Practical identifiability of ASM2d parameters - systematic selection and tuning of parameter subsets. *Water Res*. 2002;36:4113-4127.

25. Sin G, Gernaey K V, Lantz AE. Good modeling practice for PAT applications: Propagation of input uncertainty and sensitivity analysis. *Biotechnol Prog*. 2009;25:1043–1053.

26. Smyth BM, Ó Gallachóir BP, Korres NE, Murphy JD. Can we meet targets for biofuels and renewable energy in transport given the constraints imposed by policy in agriculture and energy? *J Clean Prod*. 2010;18:1671–1685.

27. Canakci M, Sanli H. Biodiesel production from various feedstocks and their effects on the fuel properties. *J Ind Microbiol Biotechnol*. 2008;35:431–441.

28. Scarlat N, Dallemand J-F, Banja M. Possible impact of 2020 bioenergy targets on European Union land use. A scenario-based assessment from national renewable energy action plans proposals. *Renew Sustain Energy Rev*. 2013;18:595–606.

29. Gog A, Roman M, Toşa M, Paizs C, Irimie FD. Biodiesel production using enzymatic transesterification – Current state and perspectives. *Renew Energy*. 2012;39:10–16.

30. Fjerbaek L, Christensen K V, Norddahl B. A review of the current state of biodiesel production using enzymatic transesterification. *Biotechnol Bioeng*. 2009;102:1298–1315.

31. Nielsen PM, Brask J, Fjerbaek L. Enzymatic biodiesel production: Technical and economical considerations. *Eur J Lipid Sci Technol*. 2008;110:692–700.

32. Al-Zuhair S. Production of biodiesel: possibilities and challenges. *Biofuels, Bioprod Biorefining*. 2007;1:57–66.

33. Tan T, Lu J, Nie K, Deng L, Wang F. Biodiesel production with immobilized lipase: A review. *Biotechnol Adv*. 2010;28:628–634.

34. Du W, Li W, Sun T, Chen X, Liu D. Perspectives for biotechnological production of biodiesel and impacts. *Appl Microbiol Biotechnol*. 2008;79:331–337.

35. Brask J, Damstrup ML, Nielsen PM, Holm HC, Maes J, Greyt W. Combining enzymatic esterification with conventional alkaline transesterification in an integrated biodiesel process. *Appl Biochem Biotechnol*. 2011;163:918–927.

36. Cesarini S, Diaz P, Nielsen PM. Exploring a new, soluble lipase for FAMEs production in water-containing systems using crude soybean oil as a feedstock. *Process Biochem*. 2013;48(3):484-487.

37. Mansouri SS, Ismail MI, Babi DK, Simasatitkul L, Huusom JK, Gani R. Systematic Sustainable Process Design and Analysis of Biodiesel Processes. *Processes*. 2013;1(2).

38. Lv D, Du W, Zhang G, Liu D. Mechanism study on NS81006-mediated methanolysis of triglyceride in oil/water biphasic system for biodiesel production. *Process Biochem*. 2010;45:446–450.

39. Pilarek M, Szewczyk KW. Kinetic model of 1,3-specific triacylglycerols alcoholysis catalyzed by lipases. *J Biotechnol*. 2007;127:736–744.

40. Cheirsilp B, H-Kittikun A, Limkatanyu S. Impact of transesterification mechanisms on the kinetic modeling of biodiesel production by immobilized lipase. *Biochem Eng J*. 2008;42:261–269.

41. Ricca E, Gabriela M, Stefano DP, Iorio G, Calabrò V, Paola M de, Curcio S. Kinetics of enzymatic trans-esterification of glycerides for biodiesel production. *Bioprocess Biosyst Eng*. 2010;33:701–710.

42. Li W, Li R, Li Q, Du W, Liu D. Acyl migration and kinetics study of 1(3)-positional specific lipase of Rhizopus oryzae-catalyzed methanolysis of triglyceride for biodiesel production. *Process Biochem*. 2010;45:1888–1893.

PART I
Background

Chapter 2: Biodiesel

In this chapter the types of feedstock, catalyst and processing options for the case study of the biodiesel process are introduced.

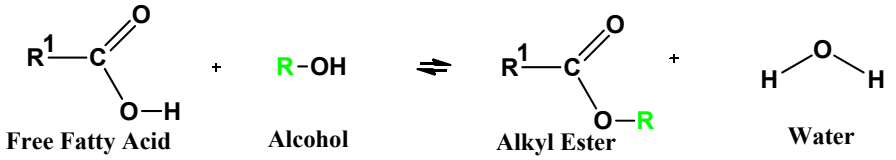
2.1. Introduction to biodiesel

Biodiesel (BD) is comprised of fatty acid alkyl monoesters derived from renewable feedstocks, such as vegetable oils, animal fats, etc. The EU has also been active in creating policy to increase the use of biofuels. In 2008 the EU adopted the Renewable Energy Directive 2009/28 (RED), which introduced a 10 percent binding target for renewable energy use in transport by 2020[1,2]. According to the European Biodiesel Board, the EU produced approximately 9.57 million metric tons of biodiesel in 2010, a 5.5 % increase compared to the previous year[3]. They also estimate that this accounts for over 40 % of the global biodiesel production and that biodiesel accounts for over 20 % of the global biofuel production.

The major operation in biodiesel production is the transesterification of the vegetable oil or animal fat into fatty acid alkyl esters (FAAE), the primary product. The main constituent of oils and fats is triglycerides (TAG), which compose about 90-98% of total mass[2]. The transesterification reaction is an ester conversion process, where the glycerol of the triglyceride, is replaced with the alkyl group of the alcohol used. Besides the transesterification reaction, there is also the esterification of the free fatty acids (FFA) found in the oil. The esterification process is a reversible reaction, where for the forward reaction, FFA are converted to fatty acid alkyl esters (FAAE) as shown diagrammatically in Figure 2-1a. The hydrolysis of alkyl esters (biodiesel) to FFA occurs in the reverse reaction. Figure 2-1b illustrates the simplified reaction scheme for the stepwise conversion of the acylglycerides to biodiesel. The transesterification using an alcohol is a sequence of three reversible consecutive steps[4,5]. In the first step, triglyceride (TAG) is converted to diglyceride (DAG). In the second step, DAG is converted to monoglyceride (MAG). In the third step, MAG is converted to glycerol. Each conversion step yields one FAAE molecule, giving a total of three FAAE molecules per TAG molecule.

This process decreases the viscosity of the vegetable oil to a value closer to that of petroleum diesel fuel while the cetane number and heating value are saved. This makes biodiesel a strong candidate to supplement petroleum diesel, as their characteristics are generally similar to that of petroleum diesel. Also since biodiesel does not contain sulphur, it does not emit sulphur oxides, and its emissions of halogens and soot are less than those of petro-diesel fuels generally used[6–8].

a) Esterification



b) Transesterification

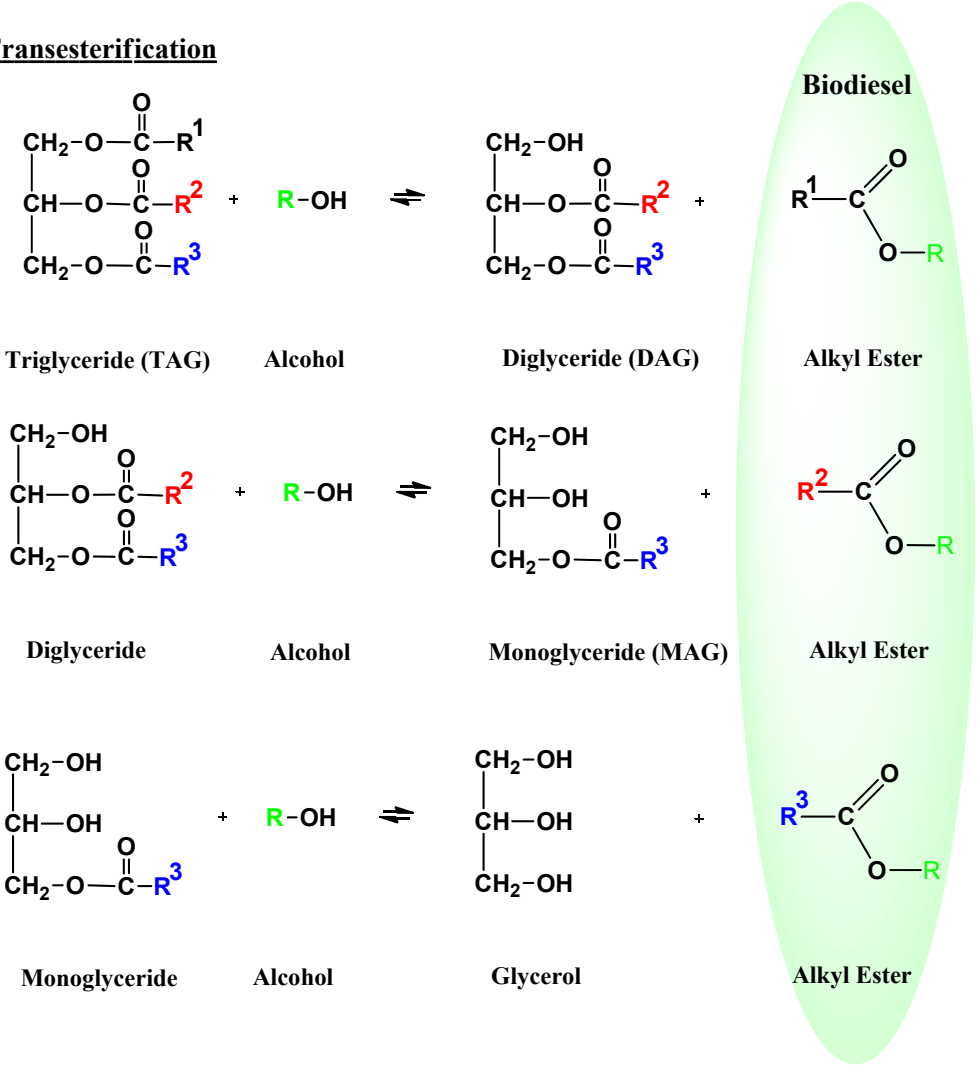


Figure 2-1 Simple depiction of the lipase (E.C.3.1.1.3) catalysed reactions in biodiesel (FAAE) production. R^1, R^2 and R^3 represent linear fatty acid chains with 12 to 24 carbon atoms which can be saturated or unsaturated. R is the alkyl group of the alcohol.

2.2. Feedstocks

2.2.1. Oils/Fats

When producing biodiesel the type of vegetable oil used varies by region. In Europe, rapeseed oil is the most commonly used oil compared to Malaysia and Indonesia where palm oil is the most significant source and soya bean oil in North America[2,5]. Fatty acids of vegetable oils vary in their carbon chain length and in the number of double bonds (Cn:x). For example the carbon chain length of Oleic acid is 18 and has 1 double bond. Some of the vegetable oils with their fatty acid compositions are given in Table 2-1. The fatty acid profile greatly determines the characteristics of the fuel produced.

Table 2-1 : Fatty acid compositions (wt %) of vegetable oils[9]

Fattyacid	(Cn:x)	Palm	Olive	Rape	Soybean	Sunflower	Sunflower	Corn
Lauric	C12:0	0.1	0.0	0.0	0.0	0.0	0.0	0.0
Myristic	C14:0	0.7	0.0	0.0	0.0	0.0	0.0	0.0
Palmitic	C16:0	36.7	11.6	4.9	11.3	6.2	4.6	6.5
Palmitoleic	C16:1	0.1	1.0	0.0	0.1	0.1	0.1	0.6
Stearic	C18:0	6.6	3.1	1.6	3.6	3.7	3.4	1.4
Oleic	C18:1	46.1	75.0	33.0	24.9	25.2	62.8	65.6
Linoleic	C18:2	8.6	7.8	20.4	53.0	63.1	27.5	25.2
Linolenic	C18:3	0.3	0.6	7.9	6.1	0.2	0.1	0.1
Arachidic	C20:0	0.4	0.3	0.0	0.3	0.3	0.3	0.1
Gadoleic	C20:1	0.2	0.0	9.3	0.3	0.2	0.0	0.1
Behenic	C22:0	0.1	0.1	0.0	0.0	0.7	0.7	0.0
Erucic	C22:1	0.0	0.0	23.0	0.3	0.1	0.0	0.1
Lignoceric	C24:0	0.1	0.5	0.0	0.1	0.2	0.3	0.1
Nervonic	C24:1	0.0	0.0	0.0	0.0	0.0	0.0	0.0

Fatty acids comprise about 94 – 96% (w/w) of the triglyceride molecule, thus it is understood that the fatty acid acyl moiety comprise most of the reactive groups in the triglyceride molecule and they greatly affect the characteristics of oils and fats[2]. Generally, biodiesel produced from oils containing a higher ratio of saturated fatty acids to unsaturated fatty acids (having one or more double bonds) may solidify and clog the fuel lines during the winter[10]. However biodiesel which are made from oils containing high levels of unsaturated fatty acids are less viscous and show higher pour and cloud points properties which makes the biodiesel more suitable for both warm and cold weather conditions. The drawback however is that the biodiesel has a lower cetane index and combustion temperature which reduces the combustion quality of biodiesel. It has been predicted that feedstocks with a high level of oleic

acid (an unsaturated fatty) are the best suited for biodiesel production given they produce a biodiesel that has characteristics similar to petroleum diesel[10].

Various plant oils and animal fats can be used to produce biodiesel; it is therefore primarily the price of the feedstock compared to the quality that decides which feedstock is being used. It has been identified that the cost of raw materials (this refers to fats and oils) accounts for more than 70% of the biodiesel production cost[11]. Thus, it is envisaged that the use of waste cooking oil should greatly reduce the cost of biodiesel. In addition, production of biodiesel from waste edible oil is considered an important step in recycling waste oil. However if the traditional way of producing biodiesel via alkaline catalyst are used the amount of water and free fatty acids in the oil is of great importance. Virgin vegetable oil and waste vegetable oil differ significantly in water and free fatty acid contents[12]. Too much water will partly hydrolyse the TAG into diglycerides (DAG), monoglycerides (MAG), glycerol and free fatty acid (FFA). The FFA will then react with the catalyst to form soaps that will facilitate the formation of emulsions, which will make the separation process difficult. The soap is furthermore binding the catalyst, so adding of extra catalyst is necessary. However, this problem can be overcome by pre-treating the oil with acid to esterify the FFA. This is where the advantage of using an enzymes catalyst is apparent, given lipases are capable of converting the free fatty acids contained in waste oils to esters[2,13].

Animal fats have also been used for the production of biodiesel. However, due to the high melting temperature, which is usually near the denaturation temperature of lipase, the reaction has to take place in an organic solvent media to dissolve the solid fat[14]. The addition of organic solvent is not recommended, as it requires the addition of a solvent recovery unit[15]. Another strategy is the recycling of the FAME phase to help solubilize the animal fats. This strategy, just like the addition of a solvent reduces the reactor capacity but avoids the separation of any solvent downstream.

Table 2-2 compares oil yields from different feedstock sources with microalgae showing significant promise. Microbial oils have significant potential given their short production cycles and can be produced by fermentation using inexpensive sources, such as CO_2 or waste water[16].

Table 2-2 Comparison of some sources of biodiesel [17]

Common Feedstock	Oil yield (L/ha)
Corn	172
Soybean	446
Canola	1190
Jatropha	1892
Oil palm	5950
Microalgae 70% oil (by wt) in biomass.	136,900
Microalgae 30% oil (by wt) in biomass.	58,700

The biggest challenges for feasible production of biodiesel from algae is the availability of an abundant, cheap and sterile CO_2 source along with the recovery of the algae oil[11,17]

As can be seen in this section, there are various oils and fats that can be used to produce biodiesel. To produce fuel grade biodiesel, the characteristics of feedstock are very important during the initial research and production stage since the fuel properties mainly depend on the feedstock properties. Hence all the work done in this thesis is with rapeseed oil; with the expectation that the experimentation and methodology used can be extended to a broader range of feedstock.

2.2.2. Alcohol

In principle all alcohols can be used in the transesterification, however methanol is by far the most used alcohol in biodiesel production. This is due to it being considerably cheaper than ethanol, and due to the greater ease of downstream recovery of unreacted alcohol[18]. However, the majority of the methanol today originates from fossil fuels sources whereas the majority of ethanol is derived from renewable sources[19]. With the increase in world ethanol production, the price of ethanol is expected to decrease which suggests that ethanol may become a competitive alternative choice of acyl acceptor[10].

2.3. Transesterification catalysts

Several aspects, including the type of catalyst, alcohol/vegetable oil molar ratio, temperature, water content and free fatty acid content have an influence on the course of transesterification. Table 2-3 gives a comparative overview of the different types of catalyst.

Table 2-3 Comparison of different technologies to produce biodiesel[20,21]

Variable	Homogeneous Alkali Catalysis	Homogeneous Acid catalysis	Solid Catalysis	Supercritical Alcohol	Lipase Catalysis	
Example catalyst	NaOH, KOH	H_2SO_4	Metal oxides, Ionic exchange resin	No catalyst	Whole-cell catalysis	porous biomass support particles
					Enzyme catalysis	Immobilized lipase
						Soluble lipase
Reaction Temp. (°C)	60–70	55–80	200–550	239–385	30–40	
FFA in raw material	Saponified products	Esters	Gasoline and lube base oils	Esters	Methyl esters	
Water in raw material	Interference with reaction	Interference with reaction	–	–	Depends on lipase used	
Yield of methyl esters	Normal	Normal	Normal	Good	Higher	
Recovery of glycerol	Difficult	Difficult	–	–	Easy	
Purification	Repeated washing	–	–	Repeated washing	None	
Cost of catalyst	Cheap	–	Medium	Cheap	Expensive	
Comment	limitation of free fatty acid content in oil/fat feedstocks.	Is applicable only to water-free oils/fats.	–	Can convert any waste oils/fats to biodieselas well as refined ones	Can esterify both FFA and TAG in one step	

2.3.1. Chemical catalysis

Transesterification reactions are conventionally alkali-catalyzed or acid-catalyzed. The catalyst used most often industrially is alkaline transesterification where raw material with a high water or free fatty acid (FFA) content needs pre-treatment with an acidic catalyst in order to esterify FFA[22]. Pre-treatment is necessary to reduce soap formation during the reaction and ease the extensive handling for separation of biodiesel and glycerol together with removal of catalyst and alkaline process wastewater (main components in the process water being glycerine, esters, soaps, inorganic acids, salts and traces of methanol). The amount of process wastewater from a traditional biodiesel plant is around 0.2 ton per ton biodiesel produced[23]. Therefore the wastewater treatment and eventual need for water reuse is a severe problem both from an energy consuming and environmental point of view[19] .

The most commonly used homogeneous alkali catalysts are sodium and potassium hydroxides, carbonates and the corresponding alkoxides. Sulfuric acid, sulfonic acid and

hydrochloric acid are usually used as homogeneous catalysts in the acid catalyzed reactions. Despite the fact that yield is very high and no soap formations, the corrosive nature of acid, very slow reaction rate and higher temperature conditions limit the use of the technology for esterification reactions[10].

Heterogeneous catalysts are categorized as solid base and solid acid. Solid catalysts have the strong potential to replace liquid catalyst as transesterification aids as they can eliminate separation, corrosion and environmental problems associated with liquid acid transesterification. Solid base catalysts include a wide group of compounds in the category of alkaline earth metal hydroxides, hydrotalcites/layered double hydroxides, alumina loaded with various compounds and alkali ion exchanged zeolites. Solid base catalysts have higher catalytic performance for transesterification than solid acid catalysts. However solid acid catalysts are preferred because of its ability to simultaneously esterify and transesterify feedstocks with a high FFA content. Heterogeneous solid acid catalysts such as resins, tungstated and sulfated zirconia, polyaniline sulfate, heteropolyacid, metal complexes, sulfated tin oxide and zeolites can simultaneously catalyze esterification and transesterification reactions; making the synthesis of biodiesel from low quality oil such as waste cooking oil containing high free fatty acids a possibility[20,24,25]. Drawbacks with using a solid acid catalyst are the lower reaction rates compared to solid base catalyst. Also, for solid acid catalyst their activity gets degraded in the presence of water as compared to using a homogeneous catalyst[26].

2.3.2. Immobilised Lipase catalysis

Compared to alkaline catalysts, enzymes do not form soaps and can esterify both FFA and TAG in one step without the need of a subsequent washing step[27]. The production of biodiesel using a biocatalyst eliminates the disadvantages of the alkali process by producing product of very high purity with less or no downstream operations[22]. Thus enzymes are an interesting prospect for industrial-scale production for reduction of production costs. This is especially the case when using feeds high in FFA. Information from literature on industrial scale enzymatic biodiesel production is quite sparse and the author has come across only two producers in china. Table 2-4 gives a summary of the plants.

Table 2-4 Industrial scale enzymatic biodiesel plants[16,28]

Company	Lvming Co. Ltd. in Shanghai, China	Hai Na Bai Chuan Co. Ltd. in Hunan Province
Stated capacity (ton /yr)	30,000	20,000
Enzyme	Immobilized lipase Candida sp. 99–125	Novozyme 435
Feed	Waste cooking oil	No info
Enzyme dosage	0.4 % enzyme to weight of oil	No info
Use of solvent	No info	No info
Reactor	Stirred tank reactor, and a centrifuge was used to separate glycerol and water	Packed bed reactor
Reported % yield of FAME	> 90% under optimal conditions	> 90% under optimal conditions

Traditionally the enzymatic production of biodiesel is manly performed using immobilized (extracellular) enzymes or immobilized whole cells (Intracellular enzymes). In both cases the enzyme is immobilized on suitable support. As the cost of lipase production is the main hurdle for commercialization of the lipase-catalyzed process[29].

It should be noted that the lipase catalysed reaction has a series of drawbacks, compared to the alkaline process. The reaction rates tend to be slower, which results in much longer reaction times. Also the high price of the enzyme makes reusability of the enzyme very important. The enzymes are furthermore very sensitive to high alcohol concentration which can cause inhibition of the enzyme. This issue however, may be remedied by the stepwise addition of alcohol while the reaction proceeds[6,20]. Also the glycerol by-product poses a potential problem as it is known to inhibit immobilized lipases, most likely by clogging of the catalyst particles. Xu and co-workers investigated the production of glycerol during ethanolysis of rapeseed oil and developed a novel dyeing method for in situ visualization of glycerol in order to study its partitioning and accumulation during the ethanolysis reaction[30]. The method developed can be used as an aid for screening supports for lipase immobilization according to their interaction with glycerol.

Nielsen and co-workers commented on the use of freeze-dried enzyme powder for biodiesel synthesis[31]. Use of such enzyme preparations need to be handled with care due to safety concerns (enzyme dust is allergenic if inhaled). An alternative is the use of stabilized liquid enzyme formulations which is cheaper than its immobilised counterpart and doesn't suffer from inhibition due to the glycerol.

2.3.3. Liquid lipase catalysis

The liquid lipase formulation is a solution of the enzyme with added stabilizers to prevent enzyme denaturation (e.g. glycerol or sorbitol) as well as preservatives to prevent microbial growth (e.g. benzoate)[31]. This makes the liquid lipase formulation significantly cheaper than their immobilized counterparts[31,32].

The ability to treat low quality feedstocks enables the use of non-edible, low-value feedstocks with high free fatty acid contents. This then shift the debate on food vs fuel and the use of land and water resources for biofuel production. Enzymatic biodiesel production becomes more attractive given a wide variety of feedstock such as waste cooking oil could be used producing fuel from a waste product. There is a strong societal need to evaluate and understand the sustainability of biofuels, especially because of the significant increases in production mandated by many countries. Sustainability will be a strong factor in the regulatory environment and investments in biofuels[33,34].

2.3.1. Non-catalytic production

Not mentioned in the table but sill an interesting area of research is the use of supercritical methanol to produce biodiesel. High yields in the order of minutes are obtained in lab scale due to the simultaneous transesterification of TAG and esterification of FFA[35]. In addition, unlike the alkali-catalyzed method, the presence of water affected positively the formation of methyl esters in a supercritical methanol method. Overall Kasteren and co-workers conclude the process can economically compete with existing conventional process but heavily depends on factors such as raw material price and plant capacity[27,36].

2.4. Process overview

The biodiesel processing steps can be divided into three main steps:

- **Pretreatment**-removal of any impurities that will affect the catalyst in the reaction step or that will disrupt the downstream processing
- **Reaction**- transesterification reaction
- **Separation**- Recovery of methanol and separation of the glycerol from the biodiesel

Pretreatment: In the production of biodiesel the quality of the feedstock greatly determines the downstream processing options. Crude and unrefined oils should be

degummed to remove impurities such as phospholipids and proteins. Shimada co-workers found that one of the main components of soybean gum are phospholipids, and the addition of 1% soybean phospholipids in refined soybean oil significantly inhibited the methanolysis[37]. Hence to improve productivity and downstream processing, degumming is essential. Common methods practiced are[38,39]:

- **Chemical refining** - The crude oil is heated to a temperature of 80–90 °C and phosphoric acid is added which serves to precipitate the non-hydratable phospholipids. The free fatty acid content is removed by initial treatment with a large excess of NaOH. The reaction between caustic soda and the free fatty acids in the degummed oil results in the formation of sodium soap, which is readily removed by a centrifugal separator. The neutralized oil is then washed with 10–20% hot water to remove traces of soap and precipitated. The oil is subsequently bleached and deodorised.
- **Physical refining** - This generally consists of a water-degumming step followed by acid degumming, neutralisation, bleaching, steam stripping to remove free fatty acids and deodorization.
- **Enzymatic treatment** - Instead of using acid degumming, one can use an enzymatic method – known as Lurgi's EnzyMax® process – which makes use of phospholipase to hydrolyse the ester bonds of the phospholipid, rendered it more water soluble, hence facilitating removal by a water wash.

For the esterification and transesterification the use of enzymes in biodiesel production can be modular. An esterification pretreatment step can be applied to an existing conventional biodiesel production in which it would be considered a retrofit. There is however the case where a new plant is built that utilizes enzymes for the entire biodiesel production.

Reaction and separation: Feedstocks high in free fatty acids cannot be converted to fatty acid alkyl esters via the conventional alkaline catalysed transesterification given the free fatty acids will react with the alkaline catalyst to form soaps (saponification reaction), reducing the biodiesel yield. In order to prevent saponification during the reaction, the free fatty acid content of the feed must be below 0.5 wt %[40]. Hence a pre-treatment step is employed where an acid catalyst (e.g. sulfuric acid) is used to convert the free fatty acids

into fatty acid alkyl esters. Alternatively enzymatic catalyst could be used for the pre-treatment step to lower the free fatty acid concentration before continuing to the alkaline transesterification as can be seen in Figure 2-2. The pre-treatment step can be carried out in both batch and continuous mode industrially. The retrofit option has the advantage of flexibility with respect to raw material quality, as well as a high quality of the glycerol by-product[41].

In Figure 2-3 the production of biodiesel using only an enzymatic catalyst is illustrated. The Main reaction is performed followed by a separation of the glycerol and further polishing of the oil phase to further reduce the free fatty acids and acylglycerides. If the biodiesel is still not within specification a final distillation step may be necessary.

The review paper by Balcão and co-workers, gives an excellent overview of the reactors used for biodiesel production via immobilized enzymes[42]. The most commonly used reactor was the batch stirred tank reactor. There is great interest in continuous production processes such as pack bed reactors (PBR) and continuous stirred tank reactors (CSTR). This is due to the fact that for large scale bulk chemical production continuous operation enables efficient use of manpower and raw materials compared to a fed-batch operation. The fed-batch operation which albeit is straightforward and an efficient means for producing biodiesel has the main disadvantage of downtime between batches. In terms of continuous biodiesel production using a liquid lipase formulation there are no reported works in the scientific literature. It is with this in mind that we look at the standard fed batch production of biodiesel using a liquid lipase formulation and investigate the continuous production of biodiesel in subsequent chapters.

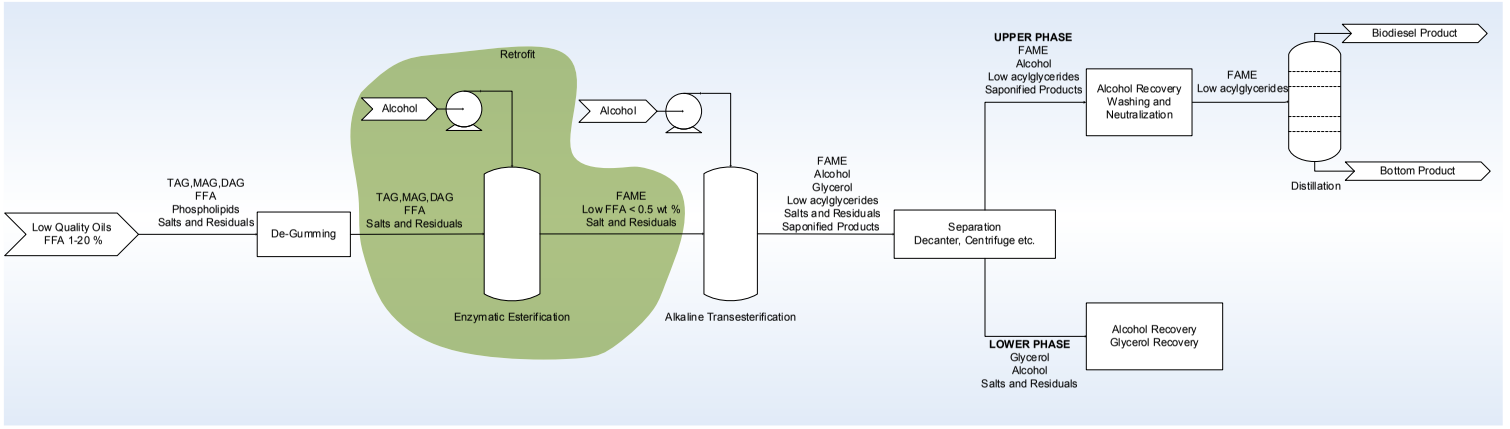


Figure 2-2 Schematic flow diagram showing the retrofit option to reduce the FFA in low quality oils before alkaline transesterification to produce biodiesel

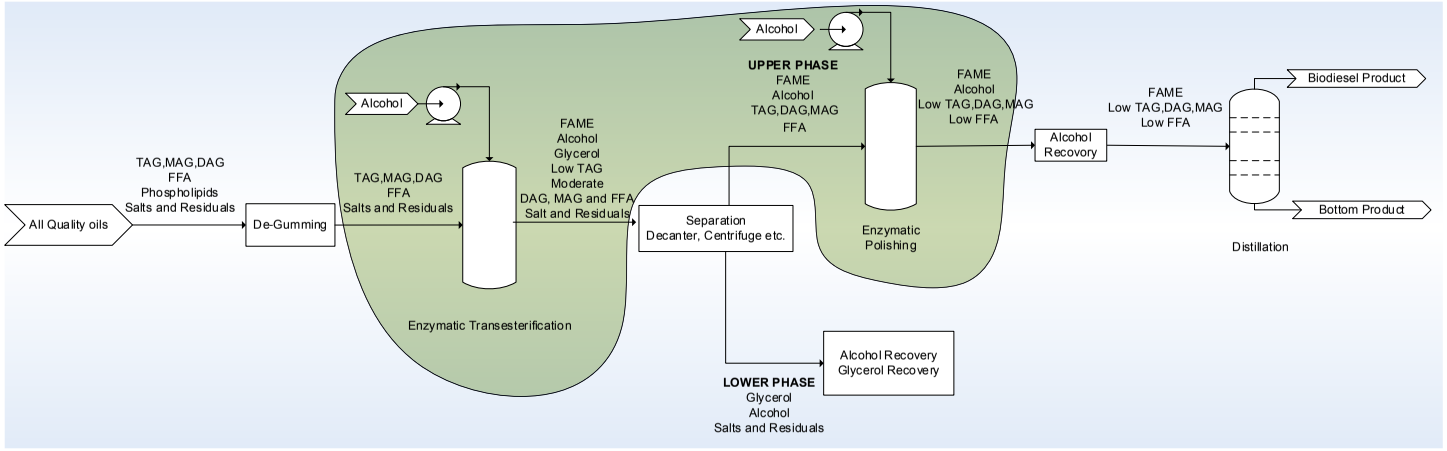


Figure 2-3 Schematic flow diagram to produce biodiesel enzymatically

2.5. Biodiesel Standards

The technical definition of biodiesel is a fuel suitable for use in compression ignition (diesel) engines that is made of fatty acid monoalkyl esters derived from biologically produced oils or fats including vegetable oils, animal fats and microalgae oils[43]. The European Standard that describes the requirements and test methods for FAME is EN 14214[41]. Biodiesel fuels can also be produced using other alcohols, for example using ethanol to produce fatty acid ethyl esters (FAEE), however these types of biodiesel are not covered by EN 14214.

Table 2-5 Condensed table of BIODIESEL standards in USA (ASTMD6751) and Europe (EN14214) [44]

Property	ASTMD6751		EN14214	
	Test method	Limits	Test method	Limits
Free glycerine	ASTMD6584	< 0.020% (w/w)	EN14105	<0.020 % (mol mol^{-1})
Total glycerine	ASTMD6584	< 0.240% (w/w)	EN14105	<0.25 % (mol mol^{-1})
Cetane number	ASTMD613	> 47	ENISO5165	>51
Cloud point	ASTMD2500	Not specified	–	Region specific
Kinematic viscosity (40 °C)	ASTMD445	1.9–6.0 mm^2s^{-1}	ENISO3104	3.5–5.0 mm^2s^{-1}

Table 2-5 is a condensed version of the European and US standards. Glycerin may be present in free or bound form (triglycerides, diglycerides and monoglycerides) and is an important parameter. Measurement of both is necessary to determine how the conversion reaction proceeded. High free glycerin content indicates poor separation and high glycerides indicate that the reaction has not proceeded to completion or the catalyst concentration is insufficient. This will lead to separation of the glycerin during storage causing, plugging of pumps and filters and can also contribute to dirty injectors or the formation of deposits on nozzles. A low flashpoint in biodiesel can result in premature ignition and can indicate residual methanol remaining from the conversion process. A fuel property that is particularly important for the low temperature operability of biodiesel fuel is the cloud point. This is defined as the lowest temperature at which wax crystals begin to form in the fuel. It is a climate-dependent requirement to allow for different seasonal grades of fuel to be set nationally. Therefore, it is an index of the lowest temperature of the fuel's usability for certain applications and the deciding factor if the biodiesel produced can be used in cold climate countries. Operating at temperatures below the cloud point of a biodiesel fuel can result in fuel filter clogging due to the wax crystals[45]. The cetane number relates to the

combustion quality of diesel fuel to self-ignite when exposed to the high temperatures and pressure in the diesel engine combustion chamber. The number is also indicative of the relative fuel stability and the EU standard requires the biodiesel to have cetane numbers of 51 or higher[46]. The establishment of these standards helps pave the way for the commercialization of biodiesel to ensure high product quality and user confidence.

2.6. References

1. Smyth BM, Ó Gallachóir BP, Korres NE, Murphy JD. Can we meet targets for biofuels and renewable energy in transport given the constraints imposed by policy in agriculture and energy? *J Clean Prod*. 2010;18:1671–1685.

2. Canakci M, Sanli H. Biodiesel production from various feedstocks and their effects on the fuel properties. *J Ind Microbiol Biotechnol*. 2008;35:431–441.

3. European Biodiesel Board. Available at: http://www.ebb-eu.org/EBBpress.php.

4. Murugesan A, Umarani C, Chinnusamy TR, Krishnan M, Subramanian R, Neduzchezhain N. Production and analysis of bio-diesel from non-edible oils—A review. *Renew Sustain Energy Rev*. 2009;13:825–834.

5. Bisen PS, Sanodiya BS, Thakur GS, Baghel RK, Prasad GBKS. Biodiesel production with special emphasis on lipase-catalyzed transesterification: Biotechnology Letters. *Biotechnol Lett*. 2010;32:1019–1030.

6. Al-Zuhair S. Production of Biodiesel by Lipase-Catalyzed Transesterification of Vegetable Oils: A Kinetics Study. *Biotechnol Prog*. 2005;21:1442–1448.

7. Parawira W. Biotechnological production of biodiesel fuel using biocatalysed transesterification: A review. *Crit Rev Biotechnol*. 2009;29:82–93.

8. Barnwal BK, Sharma MP. Prospects of biodiesel production from vegetable oils in India. *Renew Sustain Energy Rev*. 2005;9:363–378.

9. Ramos MJ, Fernández CM, Casas A, Rodríguez L, Pérez Á. Influence of fatty acid composition of raw materials on biodiesel properties. *Bioresour Technol*. 2009;100:261–268.

10. Ghaly AE, Dave D, Brooks MS, Budge S. Production of biodiesel by enzymatic transesterification: Review: American Journal of Biochemistry and Biotechnology. *Am J Biochem Biotechnol*. 2010;6:54–76.

11. Behzadi S, Farid MM. Review: examining the use of different feedstock for the production of biodiesel. *Asia-Pacific J Chem Eng*. 2007;2:480–486.

12. Al-Zuhair S, Dowaidar A, Kamal H. Dynamic modeling of biodiesel production from simulated waste cooking oil using immobilized lipase. *Biochem Eng J*. 2009;44:256–262.

13. Srivastava A, Prasad R. Triglycerides-based diesel fuels. *Renew Sustain Energy Rev*. 2000;4:111–133.

14. MEHER L, VIDYASAGAR D, NAIK S. Technical aspects of biodiesel production by transesterification—a review. *Renew Sustain Energy Rev*. 2006;10(3):248–268.

15. Al-Zuhair S. Production of biodiesel: possibilities and challenges. *Biofuels, Bioprod Biorefining*. 2007;1:57–66.

16. Tan T, Lu J, Nie K, Deng L, Wang F. Biodiesel production with immobilized lipase: A review. *Biotechnol Adv*. 2010;28:628–634.

17. Yusuf C. Biodiesel from microalgae. *Biotechnol Adv*. 2007;25:294–306.

18. Haas MJ, McAloon AJ, Yee WC, Foglia TA. A process model to estimate biodiesel production costs. *Bioresour Technol*. 2006;97:671–678.

19. Fjerbaek L, Christensen K V, Norddahl B. A review of the current state of biodiesel production using enzymatic transesterification. *Biotechnol Bioeng*. 2009;102:1298–1315.

20. Ganesan D, Rajendran A, Thangavelu V. An overview on the recent advances in the transesterification of vegetable oils for biodiesel production using chemical and biocatalysts. *Rev Environ Sci Bio/Technology*. 2009;8:367–394.

21. Watanabe Y, Shimada Y. Processes for Production of Biodiesel Fuel. In: *Biocatalysis and Biomolecular Engineering*. John Wiley & Sons, Inc; 2010:225–241.

22. Ranganathan SV, Narasimhan SL, Muthukumar K. An overview of enzymatic production of biodiesel. *Bioresour Technol*. 2008;99:3975–3981.

23. Suehara K, Kawamoto Y, Fujii E, Kohda J, Nakano Y, Yano T. Biological treatment of wastewater discharged from biodiesel fuel production plant with alkali-catalyzed transesterification. *J Biosci Bioeng*. 2005;100(4):437-42.

24. Sharma YC, Singh B, Korstad J. Advancements in solid acid catalysts for ecofriendly and economically viable synthesis of biodiesel. *Biofuels, Bioprod Biorefining*. 2011;5:69–92.

25. Helwani Z, Othman MR, Aziz N, Kim J, Fernando WJN. Solid heterogeneous catalysts for transesterification of triglycerides with methanol: A review. *Appl Catal A Gen*. 2009;363:1–10.

26. Semwal S, Arora AK, Badoni RP, Tuli DK. Biodiesel production using heterogeneous catalysts. *Bioresour Technol*. 2011;102(3):2151-61.

27. Al-Zuhair S, Ling FW, Jun LS. Proposed kinetic mechanism of the production of biodiesel from palm oil using lipase. *Process Biochem*. 2007;42:951–960.

28. Du W, Li W, Sun T, Chen X, Liu D. Perspectives for biotechnological production of biodiesel and impacts. *Appl Microbiol Biotechnol*. 2008;79:331–337.

29. Tufvesson P, Lima-Ramos J, Nordblad M, Woodley JM. Guidelines and Cost Analysis for Catalyst Production in Biocatalytic Processes. *Org Process Res Dev*. 2011;15:266-274.

30. Xu Y, Nordblad M, Nielsen PM, Brask J, Woodley JM. In situ visualization and effect of glycerol in lipase-catalyzed ethanolysis of rapeseed oil. *J Mol Catal B Enzym*. 2011;72:213–219.

31. Nielsen PM, Brask J, Fjerbaek L. Enzymatic biodiesel production: Technical and economical considerations. *Eur J Lipid Sci Technol*. 2008;110:692–700.

32. Cesarini S, Diaz P, Nielsen PM. Exploring a new, soluble lipase for FAMEs production in water-containing systems using crude soybean oil as a feedstock. *Process Biochem*. 2013;48(3):484-487.

33. Gopalakrishnan G, Negri MC, Wang M, Wu M, Snyder SW, LaFreniere L. Biofuels, Land, and Water: A Systems Approach to Sustainability: Environmental Science & Technology. *Environ SciTechnol (Environmental Sci Technol*. 2009;43:6094–6100.

34. Yang H, Zhou Y, Liu J. Land and water requirements of biofuel and implications for food supply and the environment in China: Energy Policy. *Energy Policy*. 2009;37:1876–1885.

35. Ayhan D. Biodiesel from sunflower oil in supercritical methanol with calcium oxide. *Energy Convers Manag*. 2007;48:937–941.

36. Van Kasteren JMN, Nisworo AP. A process model to estimate the cost of industrial scale biodiesel production from waste cooking oil by supercritical transesterification. *Resour Conserv Recycl*. 2007;50(4):442-458.

37. Shimada Y, Watanabe Y, Sugihara A, Tominaga Y. Enzymatic alcoholysis for biodiesel fuel production and application of the reaction to oil processing: Biofuel Production Process by Novel Biocatalysts. *J Mol Catal B Enzym*. 2002;17:133–142.

38. Clausen K. Enzymatic oil-degumming by a novel microbial phospholipase. *Eur J Lipid Sci Technol*. 2001;103:333–340.

39. Basiron Y. Palm Oil. In: *Bailey's Industrial Oil and Fat Products*. John Wiley & Sons, Inc; 2005.

40. Vyas AP, Verma JL, Subrahmanyam N. A review on FAME production processes. *Fuel*. 2010;89:1–9.

41. Brask J, Damstrup ML, Nielsen PM, Holm HC, Maes J, Greyt W. Combining enzymatic esterification with conventional alkaline transesterification in an integrated biodiesel process. *Appl Biochem Biotechnol*. 2011;163:918–927.

42. Paiva AL, Balcão VM, Malcata FX. Kinetics and mechanisms of reactions catalyzed by immobilized lipases. *Enzyme Microb Technol*. 2000;27:187–204.

43. *European Standard EN 14214. Automotive Fuels. Fatty Acid Methylesters (FAME) for Diesel Engines. Requirements and Test Methods*.; 2008.

44. Monteiro MR, Ambrozin ARP, Lião LM, Ferreira AG. Critical review on analytical methods for biodiesel characterization: 14th International Conference on Flow Injection Analysis and Related Techniques. *Talanta*. 2008;77:593–605.

45. Gui MM, Lee KT, Bhatia S. Feasibility of edible oil vs. non-edible oil vs. waste edible oil as biodiesel feedstock. *Energy*. 2008;33(11):1646-1653.

46. Erhan SZ, Dunn RO, Knothe G, Moser BR. Fuel Properties and Performance of Biodiesel. In: *Biocatalysis and Bioenergy*. John Wiley & Sons, Inc; 2008:1–57.

Chapter 3: Characteristics of Lipase Catalysis

This chapter introduces the reader to the reactions catalyzed by the lipase, the effects of the key process parameters on the enzyme catalyzed transesterification reaction and the kinetic models used in literature for biodiesel production.

Lipases (triacylglycerol acyl hydrolases, E.C. 3.1.1.3) are water-soluble enzymes and are a subclass of the esterases (An esterase is a hydrolase enzyme that splits esters into an acid and an alcohol via hydrolysis). Under the international system of classification lipases are carboxylic ester hydrolases and have been termed as glycerol-ester-hydrolases[1]. Lipases are enzymes that catalyze the hydrolysis of fats and oils with release of free fatty acids, diglycerides, monoglycerides, and glycerol[2]. Furthermore, in organic media, these enzymes also catalyse synthetic reactions including esterification, acidolysis, alcoholysis, and interesterifications as illustrated in Figure 3-1[3].

Lipases can be isolated from plants (e.g. papya latex, oat seed lipase etc.), animals (e.g. pancreatic lipase) and microorganism such as bacteria (*Pseudomonas cepacia*, *Pseudomonas fluorescens*), fungi (*Thermomyces lanuginosus*, *Rhizomucor mihei*) and yeast (*Candida rugosa*, *Candida Antarctica*)[4,5]. Microbial lipases are valued biocatalysts due to their peculiar characteristics such as the ability to utilize a wide range of substrates, high activity, stability in organic solvents, and regio- and/or enantioselectivity. These enzymes are currently being applied in a variety of biotechnological processes, including detergent preparation, cosmetics, paper production, food processing, biodiesel production, biopolymer synthesis, bio-catalytic resolution of pharmaceutical derivatives, esters, and amino acids[6].

The most commonly used lipases for enzymatic biodiesel production are those from bacteria, fungi and yeast[7]. For the biodiesel production with feeds containing both triglycerides (TAG) and free fatty acids (FFA), the employed lipases should show high activity

Hydrolysis	$RCOOR' + H_2O \rightarrow RCOOH + R'OH$
Esterification	$RCOOH + R'OH \rightarrow RCOOR' + H_2O$
Interesterification	$RCOOR' + R''COOR^* \rightarrow R''COOR' + RCOOR^*$
Alcoholysis	$RCOOR'' + R'OH \rightarrow RCOOR' + R''OH$
Acidolysis	$RCOOR' + R''COOH \rightarrow R''COOR' + RCOOH$

Figure 3-1 Lipase catalysed reactions. *R* denotes different acyl groups

on the acylglyceride substrates and FFA; as well as be non-specific so that all tri-, di- and monoacylglycerides can be converted to biodiesel. These lipases should be robust enough to tolerate moderate temperature and alcohol concentrations, exhibit low product inhibition and be able to achieve a high biodiesel yield in a short reaction time[8,9]. Many types of lipases are able to achieve the aforementioned requirements and reach conversions above 90%, in the temperature range between 35 and 50 °C. However, the optimal reaction conditions for a specific lipase are dependent on the reaction conditions, origin and formulation of the lipase. For example the reaction times to reach a given biodiesel conversion is 8 hours using Jatropha oil with immobilized lipase from *Pseudomonas cepacia* to 90 hours for the same enzyme in soluble form for the transesterification of soybean oil with methanol[8]. The following sections delve into the lipase and the effect of the process conditions on the reaction.

3.1. Interfacial activation and conformational changes of lipases

Most lipases have only a marginal activity towards molecularly dissolved substrates in aqueous solutions. However, they show high activity towards water-insoluble substrates as well as partly soluble substrates when exceeding their solubility limit, leading to the micellar aggregates or emulsions[10,11]. This unique characteristic for most of the lipases with low activity towards monomeric substrates and dramatically increasing activity above a critical aggregation concentration is the results of the phenomenon of interfacial activation[12]. In the pioneering work by Sarda and co-workers they were able to show that pancreatic lipase exhibits little activity when the water-soluble short chain triacylglycerol triacetin is in the monomeric state but the lipase activity rapidly increases when the solubility limit is exceeded as illustrated in Figure 3-2[13]. Compare this to the esterases which follows Michaelis-Menten behaviour and only act on water-soluble substrates. The same behaviour is observed for the lipase from *Thermomyces lanuginosus* (Callera Trans L utilized for this work is a modified form of this lipase). For this lipase there is a pronounced increase in activity after exceeding the solubility limit of the partially water insoluble substrate p-nitrophenyl butyrate[14] .

The increase in lipase activity is due to the structure of the lipase. Lipases share a common fold of the α/β-hydrolase type. The structure typically contains a small α-helix or

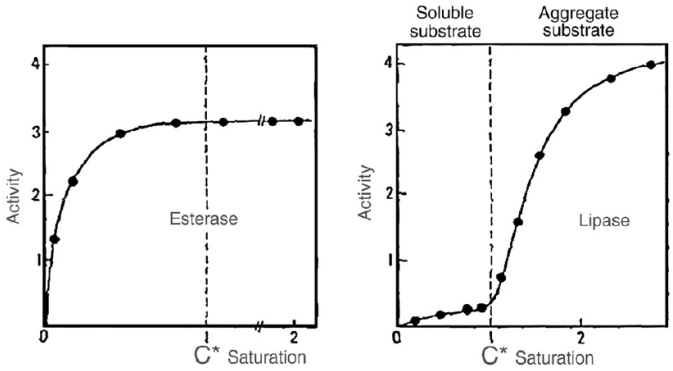


Figure 3-2 Lipase activity for the hydrolysis of triacetin by pancreatic lipase (right panel) and an esterase (left panel) as a function of the substrate concentration of the partly water-soluble ester. The dashed lines represent the limit of solubility of the used ester[13].

loop acting as lid that covers in the closed (inactive) conformation the active site pocket. Thus the active site is not accessible for the substrates leading consequently to none or limited conversion.

When the lipase is exposed to a lipid/water interface, the lid is displaced and undergoes a conformational rearrangement to the open (active) conformation so that the active site pocket becomes accessible for the substrates[12,15] .

Due to an opposite polarity between the enzyme (hydrophilic) and their substrates (lipophilic), lipase reaction occurs at the interface between the aqueous and the oil phases. The interfacial activation due to the conformational rearrangement of the lid is the result of a changing surrounding environment around the enzyme. The adsorption at liquid/liquid interfaces enables the more hydrophobic unpolar parts to protrude into the non-polar phase leading to a decrease in Gibbs free energy. The extent of binding to emulsified or aggregated substrates is related to physicochemical properties and the compositional structure of the interface that is often described in literature as “quality of the interface”[11]. It has also been reported that substrates can modify the quality of the interface. For example, sn-2 monoglycerides (2-position ester group on the glycerol backbone) tend to occupy the interface and expel free fatty acids, diglycerides, triglycerides and sn-1,3 regiospecific lipases from the oil–water interface[11]. Hence, interfaces are the key spots for lipase biocatalysis and it is not always straightforward to differentiate between substrate inaccessibility and enzyme denaturation/inactivation[11].

It should be noted that not all lipases exhibit the phenomenon of interfacial activation. The interfacial activation of *Candida Antarctica* lipase A (CALA) and B (CALB) was compared to that of the *Thermomyces lanuginosus* lipase (TLL) (former known as *Humicola lanuginosus* lipase)[14]. CALB displayed no interfacial activation while CALA displayed a marginal interfacial activation when using partially soluble p-nitrophenyl butyrate (pNPB) and increasing the concentration. However, TLL compared to CALB and CALA displayed pronounced interfacial activation.

3.2. Reaction mechanism

Lipases catalyze reactions where two substrates react to two products. The reaction mechanism for lipase-catalyzed reactions such as esterification of long-chain fatty acids as well as hydrolysis and transesterification of glycerol esters follows a Ping-Pong Bi-Bi mechanism [16,17], as illustrated schematically in Figure 3-3.

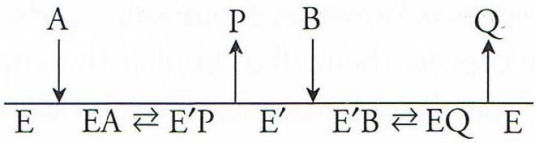


Figure 3-3: Ping-Pong Bi Bi or substituted-enzyme mechanism, respectively[18].

The first substrate (A) binds to the enzyme (E) and forms the first enzyme substrate complex (EA). The first product (P) is released after which the acylated enzyme complex (E') binds to the second substrate (B) to form the second enzyme substrate complex (E'B).The second substrate is then released along with the enzyme.

A lipase is able to generate a nucleophilic residue for covalent catalysis by using an Acid-Base-Nucleophile triad. The catalytic triad is composed of the amino acids serine, histidine and aspartate or glutamate. The residues form a charge-relay network to polarise and activate the nucleophile, which attacks the substrate, forming a covalent intermediate which is then hydrolysed to regenerate free enzyme. The nucleophile in this case is the serine molecule. With this in mind a more in depth mechanism for the enzymatic production of FAME (biodiesel) from triacylglycerides based on a classical Ping-Pong Bi Bi mechanism is proposed by Al-Zuhair and co-workers and is illustrated in Figure 3-4[19].

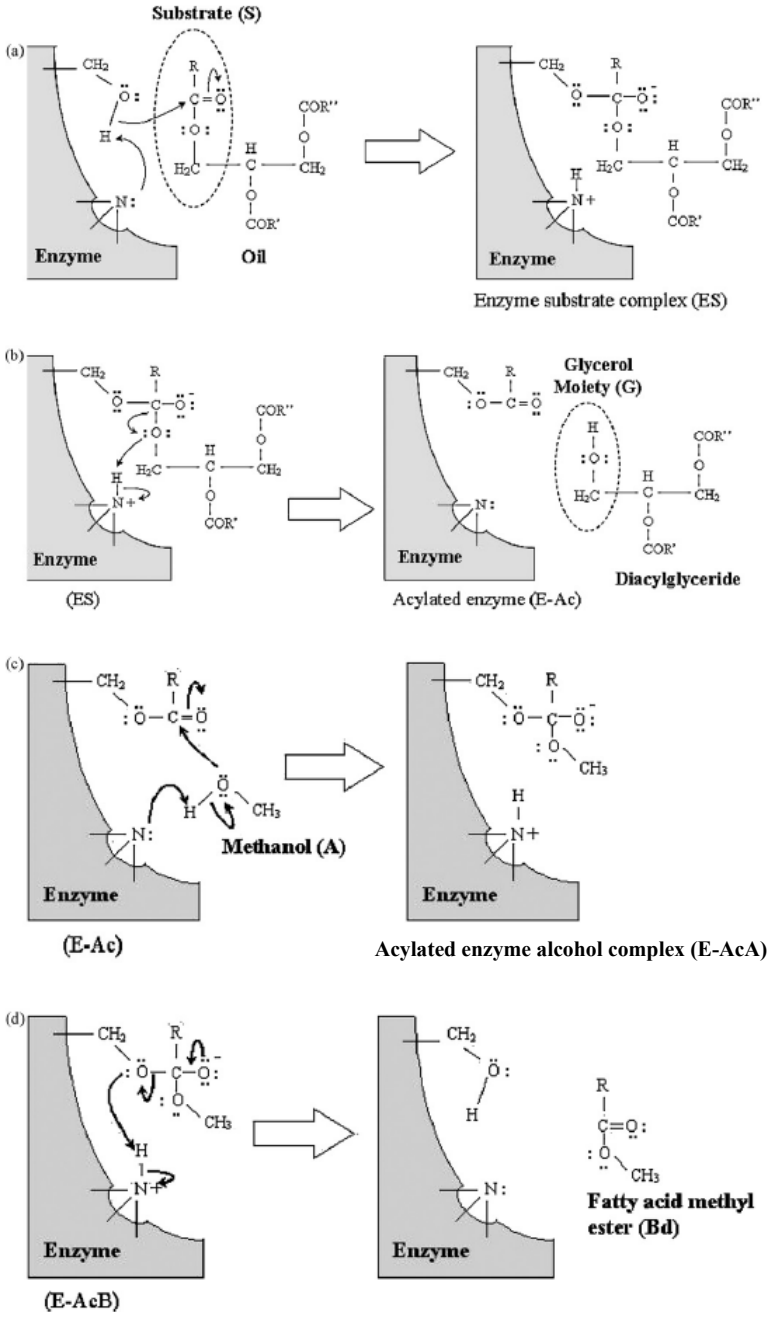


Figure 3-4: The catalytic mechanism of the lipase-mediated transesterification of triacylglycerides[19].

The first part of the catalytic mechanism is the cleavage of the ester bond and the formation of the acyl-enzyme intermediate. In the first step (a), the nucleophile (serine) is made more active when the base in a catalytic triad (histidine) polarises and deprotonates the nucleophile to increase its reactivity and the acidic residue (aspartate or glutamate) stabilise the deprotonated state during the catalytic cycle. Consequently, the hydroxyl group of serine is able to attack the carbon atom of the carbonyl group forming a tetrahedral intermediate. The negative charge of the tetrahedral intermediate is stabilized by the peptidic NH groups of the oxyanion hole contributing to charge distribution and reduction of the ground state energy of the tetrahedral intermediate. Afterwards (b) a proton is transferred from the conjugated acid of the histidine to the alkyl oxygen atom of the bound substrate. This leads to the cleavage of the substrate ester bound releasing the glycerol moiety and the formation of the acyl-enzyme intermediate. If a TAG was the initial substrate, then a DAG would be formed, whereas a MAG would be formed in case of DAG as initial substrate. The second main part is the esterification of the covalent bound acyl group with an alcohol (in case of hydrolysis with water) in order to form and release FAME (or free fatty acid in case of hydrolysis). During the deacylation steps the hydroxyl group of the alcohol is firstly deprotonated by the histidine in order to enhance the nucleophilicity (c). Subsequently, the carbon atom of the carbonyl group bound to the enzyme is attacked, which forms a second tetrahedral intermediate stabilized by the oxyanion hole. Finally (d), the serin oxygen atom is protonated by the protonated histidine resulting in the release of the final product biodiesel and the regeneration of the catalytic site [19].

When working with water-oil systems as in this study the reactants are localized in different phases as the lipids are heterogeneous molecules being mostly insoluble or partly soluble in water and forming mostly micelles or emulsions[10].

The lipase Callera Trans L was found to be an interfacial activated enzyme that needs to adsorb at an interface in order to be activated and reactive. Verger and co-workers proposed a simplified two-dimensional Michaelis-Menten mechanism for the interfacial hydrolysis of phospholipids by porcine pancreatic phospholipase A[20]. The mechanism is depicted in Figure 3-5 and illustrates the enzymatic reaction occurring at the interface (hatched area) and not in the bulk phase.

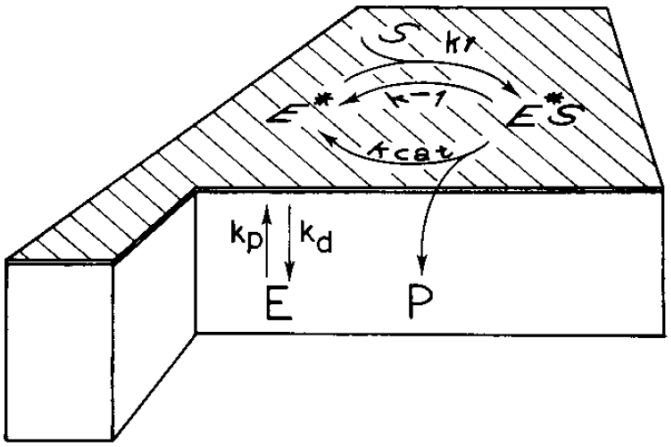


Figure 3-5: Proposed model for the action of a soluble enzyme at an interface[20].

The first step is the reversible adsorption of a water-soluble enzyme at an interface. This step is also described as penetration into an interface leading to a more favourable energetical state denoted as E^*. After the adsorption step the enzyme at the interface binds to the substrate molecule (*S*) at the interface resulting in an interfacial enzyme substrate complex (E*S). The catalytic step then follows where the penetrated enzyme is regenerated at the interface E^*, releasing the product in the water phase.

3.3. Substrate Specificity of Lipases

The type of substrate will affect the catalytic rate. The specificity of the lipases is due to the regiospecificity and specificity with respect to the length of hydrocarbon chain of fatty acid[21]. Two main factors that determine the fatty acid and alcohol selectivity of lipases are the inductive effect and steric hindrance[22].

3.3.1. Inductive effect

It is postulated that there is a nucleophilic attack of the hydroxyl residue of the lipase on the carbonyl group of the substrate. The inductive effect then explains the decrease in rate of reaction for various substrates. For example, as the carbon that is attacked is more electrophilic, the rate of reaction is faster e.g. triglyceride > 1,2-diglyceride > 1,3-diglyceride > 1-monoglyceride > 2-monoglyceride[22].

3.3.2. Steric hindrance

The bulkiness of the carbonyl group inhibits the rate of reaction. For example vinyl oleate is hydrolysed, but phenyl oleate is not. This steric effect also explains the specificity of the lipase for the α-chains of triglycerides.

Lipases can be divided according to their regioselectivity with regards to the acyl chains of the triglycerides; sn-1,3 regiospecific (e.g. lipase from *Rhizomucor miehei* and pancreatic lipase) or non-regiospecific (e.g. lipase from *Candida rugosa*)[23]. Regiospecificity is referred to the acyl position on the glycerol backbone that is preferred to be cleaved by a regiospecific lipase as depicted in Figure 3-6.

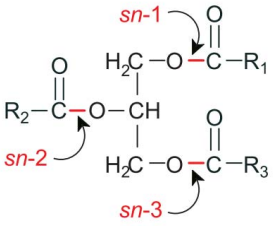


Figure 3-6: The *sn*-1, *sn*-2 and *sn*-3 positions in triacylglycerides.

The specificity towards ester bonds in positions *sn*-1,3 of the triacylglycerides may result from the inability of lipases to act on the *sn*-2 position of the triacylglycerol. Steric hindrance was reported to prevent the entrance of the acyl group in *sn*-2 position to the active site. A transesterification reaction employing a *sn*-1,3-specific lipase such as TLL may initially produce a mixture of 1,2-, 2,3-diacylglycerols and FAME as products. However, *sn*-1,3-specific lipases can achieve surprisingly 90% biodiesel yield exceeding the maximum theoretical yield of 66%[21]. There are two approaches in literature explaining how the formed 2-MAG could be converted. Hermansyah and co-workers proposed that *sn*-1,3-specific lipases can cleave both ester bonds in *sn*-1- and *sn*-2-position but with different rates. It was assumed that the cleavage of the ester bonds at the extreme positions proceeds faster than the cleavage of the ester bond at *sn*-2-position. Consequently, both 1-MAG and 2-MAG are formed, whereby the fraction of 2-MAG is greater than the fraction of 1-MAG. 2-MAG can be converted by *sn*-1,3-specific lipases but more slowly than 1-MAG[24]. Pilarek and co-workers postulated that only one form of DAG (1,2-DAG) and only one form of MAG (2-MAG) are formed by *sn*-1,3-specific lipases. The formed 2-MAG can then be

converted to glycerol by acyl migration of the acyl group from the *sn*-2 position to either the terminal *sn*-1 position or *sn*-3 position forming 1(3)-monoacylglycerols as can be seen in Figure 3-7, before 1-MAG can be converted by the *sn*-1,3-specific lipases[25].

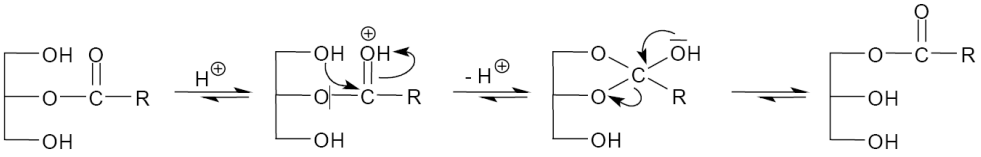


Figure 3-7: Non-enzymatic acyl migration in 2-monoacylglycerides[26]

The s*n*-1 and s*n*-3 positions in triacylglycerols are sterically distinct. Some lipases are able to differentiate between the two primary esters at the s*n*-1 and s*n*-3 position. Lipases from *Pseudomonas* species and the porcine pancreatic lipase have shown stereoselectivity when certain acyl groups are hydrolyzed[23].

3.4. Key parameters affecting the enzymatic biodiesel reaction

3.4.1. Impact of the temperature

The Arrhenius' equation can be used to estimate the temperature dependence of reaction rates. However, in the case of enzyme-catalysed reactions one has to consider the temperature denaturation of the enzyme. Denaturation involves the rupture of intramolecular interactions such as hydrogen bonds and other weak interactions leading to an unfolding of the enzyme and consequently to a loss of catalytic activity. With this being said, in an assay, enzyme catalysed reactions exhibit a pronounced temperature optimum for a given set of conditions[18].

Optimum temperatures for enzymatic biodiesel production from triacylglycerides and fatty acids in solvent-free systems and systems with n-hexane as solvent varied between 30 and 50 °C. Dizge and co-workers studied the effect of the temperature on the catalytic activity of immobilized TLL used for transesterification of canola oil with methanol in the range of 30 to 70 °C. For temperatures above 50 °C, the enzyme lost its activity dramatically[27]. 40 °C was found to be the optimal temperature with the formation of 85.8% FAME as final yield. Chen and co-workers reported a decrease of both the initial reaction rate and the final yield with temperatures higher than 35 °C for the NS81020-mediated esterification of oleic acid with methanol[28]. Al-Zuhair and co-workers observed an increase

of the apparent rate constant up to 50 °C after which the initial reaction rate decreased sharply after 50 °C due to thermal denaturation in case of the employed *Candida rugosa* lipase[29]. It should be noted that the optimal temperature of the lipase preparation can be expected to increase when immobilizing the lipase on a carrier. The binding to the support stabilizes the lipase and decreases the effect of thermal denaturation as the binding decreases the degree of freedom for the unfolding of the protein structure. This leads to less increase of entropy, which stabilises the native form of the enzyme[8,9].

Al-Zuhair and co-workers have shown that the increase of the reaction rate could also be partly attributed to an increase of interfacial area with increasing temperature[29]. With increasing temperature the average droplet diameter and viscosity decreases. This results in a decreasing surface tension leading to smaller droplets. Consequently the temperature affects both the emulsion and the enzyme.

3.4.2. Impact of the pH

The pH affects both the activity and the stability of the lipase. The dependency of the lipase activity on the pH is mainly caused by their origin and the dependency of the ionization state of the catalytic active amino acids at the active site[23]. The serine residue of the catalytic triad is only in a deprotonated form active. As already explained in section 3.2, the first step of the catalytic mechanism is that serine is activated by deprotonation, for which histidine and an acid amino acid (aspartate or glutamate) are required. Consequently, the imidazol ring of the histidine and the carboxyl group of the acid residue have to be present in a deprotonated form in order to enable the deprotonation of the serine hydroxyl group.

The pH optima for most of the lipases lie in the range between pH 7 and 9[23]. Chen and co-workers observed in the case of esterification of oleic acid with methanol using the soluble lipase NS81020 showed an optimal pH of 6.86[28]. However, the highest biodiesel yield was found to be at pH = 10.55. They postulated this might be due to the hydroxyl ions suppressing the hydrolysis of the formed oleic acid methyl ester (biodiesel) and a facilitation of the esterification by methanol.

Similar to the temperature dependency, the pH optimum can be affected by immobilization as well. The shift of the pH optimum when immobilizing the enzyme on a

support might be mainly due to the partitioning of protons between the bulk phase and the microenvironment around the support. The pH of the microenvironment around the immobilized enzyme may be also affected by the acid-base properties of the support[23]. Thus the pH-optimum of the same lipase in immobilized form could be different compared to the pH-optimum for the free lipase.

The influence of the pH on the enzyme stability can be explained by the net charge of the protein surface, which is affected by the pH of the surrounding solution. If the pH equals the isoelctric point, the enzyme carries no electrical charge on its surface. Consequently, less water molecules can be taken up by the protein surface decreasing the solubility of the enzyme. This could lead to precipitation of the enzyme leading to less active enzyme molecules that can catalyze the biodiesel production[30].

3.4.3. Impact of the water content

The water content in the reaction mixture is one of the most important factors in the lipase-catalyzed production of biodiesel as it has a great impact on both the activity of the lipase and the thermodynamic equilibrium and consequently on the final biodiesel yield. A minimum amount of water surrounding the lipase is essential to maintain the three dimensional conformation of the lipase, especially if used in a soluble form[8,31]. In case of too low water content the lipase would be denaturated, whereby the catalytic activity is lost. Callera Trans L also requires an oil-water interface in order to be active towards the glycerol esters or free fatty acids. Increased water content increases the volume of the dispersed polar phase leading to an increased interfacial area. On the other hand an excess of water promotes the hydrolysis reaction. Therefore, more free fatty acids (FFA) are formed in case of higher water contents. It is admittedly possible to esterify these formed free fatty acids but additional water will be formed as a by-product of the esterification. According to Le Chatelier's principle, this would shift the thermodynamic equilibrium producing lower biodiesel yields. Besides, excessive water reduces the alcohol concentration. This leads to a lower probability for a nucleophilic attack on the carbonyl carbon atom of the acyl-enzyme intermediate of methanol, which suppresses the esterification of fatty acids and transesterification of glycerol esters resulting in a slower reaction rate[9,28]. Consequently, the optimal water content is a compromise between minimizing the hydrolysis and maximizing

the amount of activated enzyme molecules at the interface for a given enzyme concentration.

The optimal water content depends additionally on the source of the lipase. As shown by Kaieda and co-workers the optimal water content for the transesterification of soybean oil with methanol by soluble lipases from *Candida rugosa*, *Pseudomonas cepacia* and *Pseudomonas fluorescencs* varied from lipase to lipase[32]. The formation rates of the fatty acid methyl esters by the lipases from *C. rugosa* and *P. fluorescencs* initially increased with increasing water content and decreased after reaching a certain optimum. In both cases, the reaction rates significantly decreased when the water content was too low. The final yield as well as the formation rate of biodiesel was additionally not significantly improved by increasing the water content. But all investigated lipases commonly exhibited no activity, if the system was water-free. Consequently, soluble lipases especially require an essential amount of water in order to exist in their native active conformation. Cesarini and co-workers investigated the biodiesel production with crude soybean oil and methanol as substrates by the soluble TLL lipase, Callera Trans L[31]. They observed significant lower biodiesel production of 88.2% at a water content of 15% (w/w of oil) compared to over 95% biodiesel yield in case of 3 and 5% (w/w of oil) water accompanied by an increase of free fatty acids with an increasing water content. Obviously, the higher water content shifted the biodiesel reaction to lower final biodiesel yields.

3.4.4. Impact of the methanol concentration

Stoichiometrically, three molar equivalents of alcohol are required for a complete conversion of triglycerides in the oil and at least one molar equivalent of alcohol is required for a full conversion of the free fatty acids. However, transesterification and esterification are both reversible reactions. An increase in the amount of the alcohol as one of the reactants would consequently drive the thermodynamic equilibrium toward the product site and increase the biodiesel yield.

However, excessive alcohol can cause both reversible competitive inhibition[9,19] and irreversible inactivation of the lipase due to a destabilizing effect of especially short-chain alcohols such as methanol on the stability of the lipase[21,27,33].

The alteration of enzyme activity due to inactivation in the presence of polar organic solvents such as alcohols can be the result of the following phenomena. Organic solvents such as short-chain alcohols tend to strip off water molecules from the hydration shells of proteins. Consequently the distance between lipase molecules decreases as a result of a thinner hydration shell. The lipases proteins approach each other and might precipitate due to intermolecular interactions such as van der Waals interactions[30].

Chen and co-workers concluded from their experimental results for the esterification of oleic acid catalyzed by the soluble lipase NS81020, that both the presence of water and methanol addition strategy contributed to the inhibiting and inactivating action of methanol on the lipase activity[28]. A lower methanol concentration should be preferred in order to enable reuse of the enzyme formulation. Since an alcohol excess is prerequisite of high biodiesel yields, an optimal alcohol to oil (or free fatty acid) ratio at the beginning and an optimal addition strategy of the alcohol is required in order to circumvent alcohol inhibition and inactivation of the enzyme. A strategy with adding alcohol in successive steps has been for instance applied by Cesarini and co-workers[31]. They achieved 96% biodiesel yield after 24 hours using a stepwise methanol addition strategy.

3.4.5. Impact of the enzyme concentration

In general, the initial reaction rate increases with increasing lipase concentration until a certain concentration at which the initial reaction rate remains constant even after adding more enzyme[9]. For a soluble lipase it was observed that there was a linear increase of the initial reaction rate with increasing enzyme concentration in the bulk phase for low enzyme concentrations, followed by levelling off and reaching a constant value similar to a Langmuir-isotherm[28,29]. Free lipases are adsorbed and desorbed continuously at the oil-water interface[34]. In this dynamic system the interfacial area is partly covered with adsorbed enzyme at any instant of time. Initially, the interfacial enzyme concentration increases linearly with increasing enzyme concentration in the bulk phase as there are enough free surface places for the adsorption of the lipase. With further increasing enzyme concentrations in the bulk phase, it is more difficult for the lipases to adsorb at the interface until all surface places are occupied by penetrated lipase molecules reaching a maximal surface concentration. Hence, although an increase of the enzyme concentration in the bulk phase is assumed to increase the interfacial concentration and consequently the reaction

rate, there is a maximal enzyme concentration at which the interface becomes saturated. Beyond this point, any increase in the enzyme concentration in the bulk phase would not enhance the reaction rate.

3.4.6. Impact of the mixing

The reaction system in case of enzymatic biodiesel production is for both immobilized lipase formulations and soluble lipase formulations a multi-phase reaction system. Consequently, mass transport phenomena have a great impact on the reaction rate and on the productivity of the process. In the case of a biphasic water-oil system, when employing soluble lipases as catalyst for the biodiesel production, the mixing affects the mass transfer along with the emulsification process.

For an oil water emulsion, increasing stirring speed causes an increase in the interfacial area as a result of a decrease in the average droplet diameter with increasing stirring speed[28,35].

The volumetric power density (w/L) represents the power input per unit volume. The supplied power of the agitator is a function of the rotational speed of the agitator. With increasing stirring speed the supplied volumetric power input increases resulting in a decreased maximal average drop diameter due to an increased shear rate on the drops that splits the larger drops into smaller drops. The total specific interfacial area a_t, as defined in equation (3.1) with d_{mean} representing the sauter mean diameter of the droplets and φ representing the volume fraction of the emulsified phase[35].

$$a_t = \frac{6}{d_{mean}} \cdot \varphi \quad \text{Eqn (3.1)}$$

With increased interfacial area there is a greater potential for the formation of the penetrated lipase at the interface (see Figure 3-5). Due to more of the penetrated lipases at the interface more substrate molecules can be converted resulting in higher initial reaction rates.

3.5. Transesterification Models

The key parameters affecting the enzymatic biodiesel reaction have been presented and now we investigate how these phenomena are modelled in the scientific literature. In order

to model the enzymatic biodiesel reaction, it is important to elucidate the various phenomena that need to be considered when modelling the enzymatic biodiesel system.

Descriptions of the various kinetic models for enzymatic transesterification of vegetable oils are quite numerous[36,37,25,38–41]. Table 3-1 gives an overview of some of the most recent kinetic models in literature for the transesterification reaction and highlights the various phenomena/mechanisms that each particular model addresses.

3.5.1. Various concentrations and Water in oil emulsions

The most rigorous kinetic model considered the difference between the interfacial and bulk concentrations of the enzyme, substrates and products. This is important given Lipases occur in alternative conformational states stabilised by the interaction with the water/substrate interface. In order to describe the differences between the interfacial and bulk concentrations of the enzyme, substrates and products, linear/nonlinear relationship were incorporated. Normally a nonlinear relation, such as the Langmuir adsorption model, is introduced. On the other hand, for the substrates and/or products, the linear relationships between the interfacial and bulk concentrations were usually incorporated or it's assumed the interfacial concentrations are the same as in the bulk reaction system.

3.5.1. Mixing/Power Input

Only Al-Zuhair and co-workers in their earlier work investigated how the interfacial area of the oil-water emulsion varied with rotational speed of the mixer for the hydrolysis of palm oil [29]. They have formulated various correlations which are specific to the system they are working with given it is an empirical correlation. However, it can still be used to give an order of magnitude estimate of the interfacial area available. It should be noted that in Al-Zuhair and co-workers work the correlations were based on an oil in water system where water constituted the bulk phase (over 60 v/v %)[29].

3.5.2. Temperature

Given that the reaction is usually carried out at a fixed temperature in the range of 30-45 ^{o}C, most of the kinetic models proposed ignore how the rate of reaction varies with temperature. This is reasonable given during the reaction a fixed temperature is used.

Table 3-1 Recent kinetic models describing enzymatic transesterification

		Pilarek (2007)[25]	Al-Zuhair (2007)[19]	Cheirsilp (2008)[38]	Dan Lv (2008)[36]	Hermansyah (2010)[24]	Fedosov(2013)[41]	General Comments
Type of Reaction		Transesterification	Transesterification	Transesterification	Transesterification	Hydrolysis	Transesterification	
Substrate		Triacetin or Triaprylin and 1- or 2-Propanol	Palm oil and Methanol	Palm oil and Ethanol	Soyabean oil and Methanol	Triolein and Oleic acid	Rapeseed oil and Methanol	
Enzyme	**Type Enzyme**	CaLB sn-1,3 specific Lipase Immob. Macro porous acrylic resin. Novozym 435	Mucor Miehei, sn-1,3 specific Lipase, solubilized Lipase	Pseudomonas sp, sn-1,3 specific Lipase, Immob. Microporous polypropylene powder	NS81006 from Aspergillus niger, sn-1,3 specific Lipase, solubilized Lipase	Suspended Porcine pancreatic Lipase, sn-1,3 specific Lipase	CaLB sn-1,3 specific Lipase Novozym 435	Enzyme formulation will dictate how the various phases in the system should be handled
	Treatment of Enzyme concentration	Enzyme concentration treated as a bulk concentration	Enzyme concentration treated as a bulk concentration	Enzyme concentration treated as a bulk concentration	No Consideration. Fixed amount of Enzyme used	Effect of the interfacial area available for the enzyme is considered.	Enzyme concentration treated as a bulk concentration. Also introduces various pseudo phases	Interfacial area needs to be considered for lipases that are activated by the oil water interfacial area
	Alcohol Inhibition	Competitive	Competitive	Competitive	No Consideration	Not applicable	Competitive	Also seen in our work and needs to be considered
	Acyl migration or Regio-Specificity of enzyme	Described as a first order reversible reaction	No Consideration	No Consideration	No Consideration	Accounts for positional specificity of the Lipase	No Consideration	Most likely accounts for accumulation of DAG
	Enzyme deactivation by alcohol	First order in enzyme concentration first order in alcohol concentration	No Consideration	No Consideration	No Consideration	Not applicable	First order in enzyme concentration first order in alcohol concentration	Also seen in our work and needs to be considered

		Pilarek (2007)[25]	Al-Zuhair (2007)[19]	Cheirsilp (2008)[38]	Dan Lv (2008)[36]	Hermansyah (2010)[24]	Fedosov(2013)[41]	General Comments
Various Temp.		No consideration	No Consideration	No Consideration	No Consideration	No Consideration	Introduces average temperature coeff.	Usually operate at Fixed temperature
Various Oil Compositions with no solvent		Possibly	No	Possibly	Possibly	Possibly	Possibly	Different types of oil feeds mean recalibrating parameters. Hence work with a specific feed
Mixing/Power Input		No consideration	No consideration	No consideration	No consideration	No consideration	No consideration	Usually operate at a fixed power input
Types and Scale of reactors		Batch reaction, 25-50 mL flask, water bath shaker	Batch Reaction, 1.5L reactor, 2 rushton impellers	Batch Reaction,(5 g palm oil used)	Fed Batch Reaction, Methanol dosed at 1,2,3 hrs, 250 mL reactor, mechanical stirring	Batch Reaction, cross-sectional area of reaction vessel varied to vary interfacial area	Batch reaction, 50 mL flask, shake incubator	Standard practice to combine rate equation with reactor mass balance
Types of equations	**Type of Mechanism**	Ping-Pong Bi-Bi	Ping-Pong Bi-Bi	Ping-Pong Bi-Bi	Second Order reversible reaction	Ping-Pong Bi-Bi	Ping-Pong Bi-Bi	General consensus Ping-Pong Bi-Bi
	Mathematical expression	Mass balance coupled to rate Eqn. Differential Eqns.	Initial Rate Equation	Mass balance coupled to rate Eqn. Differential Eqns.	Differential Eqns.	Mass balance coupled to rate Eqn. Differential Eqns.	Mass balance coupled to rate Eqn. Differential Eqns.	Mass balance coupled to rate Eqn. Differential Eqns.

It means however that kinetic parameters determined at one temperature need to re-estimated if operating at another temperature.

3.5.3. Types of Reactors

All the kinetic models proposed have been either for batch/fed-batch and have not been extended using the mass balance for continuous reaction systems.

3.5.4. Alcohol Inhibition and enzyme deactivation

Most authors include the competitive alcohol inhibition which is modeled as being reversible. Methanol is poorly miscible with oil/fat and tends to deactivate the enzymes at high concentrations and is modeled as being irreversible. To overcome the deactivation of the enzyme, multi-step addition of methanol is proposed[42]. The challenge is being able to distinguish between the two phenomena when fitting experimental data.

3.5.5. Types of Equations used

The general consensus is that the transesterification reaction proceeds via a Ping-Pong Bi-Bi mechanism. The free enzyme (E) reacting with triacylglycerol (T) to form the first complex (E·T)

T is then hydrolyzed to diacylglycerol (D) and fatty acid (F). Subsequently, D is released from the second complex (E·D·F) to form the third complex (E·F). This complex might react with alcohol (Al) through an alcoholysis reaction to form an alkyl ester (Es) or with water (W) through a hydrolysis reaction to form FFA (F). The mechanism for the hydrolysis of D and monoacylglycerol (M) is also similar to that described above and is illustrated in Figure 3-8.

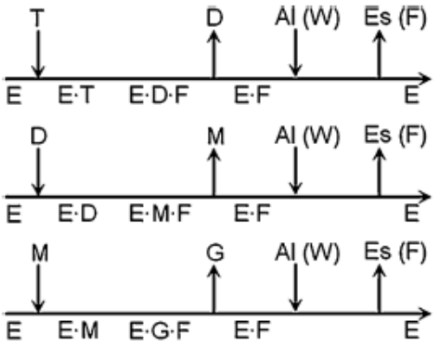


Figure 3-8 Illustration of the Ping-Pong Bi-Bi mechanism for stepwise transesterification of acylglicerides[38]

Where authors differ in the formulation of their differential equations is in the assumptions made, mainly:

- How the oil water interfacial area is handled
- The rate determining step
- Steady-state approximation for the enzyme and its complexes

The main points of contention are the last two points. In the traditional approach the steady-state approximation and the rapid equilibrium approximation are used for the enzyme complexes. The use of mass action kinetics (elementary rates of reaction) are preferable for complex kinetic schemes where networks of coupled ordinary differential equations can be used to describe the kinetics in a mass action approximation [43]. Hence we move away for making assumptions where the rate-determining step is not immediately apparent or for systems that involve multiple-tight binding interactions.

3.6. Conclusions

The main phenomena affecting the transesterification reaction have been presented. This is done to highlight what characteristics the kinetics that is to be coupled to the mass balance of the system needs to have to be used in model based design framework. This info is used in the kinetic model development in Chapter 6 to simulate and fit data over a wider range of experimental conditions (variation in water, enzyme and methanol loading) and time ranges. You may ask given the numerous kinetic models that have been presented in this chapter, why is there a need for another one? What is noticed is that the identifiability and uncertainty in the kinetic model parameter estimates are not reported which then affects the trust that can be placed in the model predictive capability. This is brought out in Chapter 5 were the uncertainty in the model predictions are investigated for one of the more promising kinetic models.

3.7. References

1. Malcata FX. *Engineering Of/with Lipases*. Kluwer Academic Publishers; 1996.
2. Svendsen A. Lipase protein engineering. *Biochim Biophys Acta*. 2000;1543(2):223-238.
3. Sharma R, Chisti Y, Banerjee UC. Production, purification, characterization, and applications of lipases. *Biotechnol Adv*. 2001;19(8):627-662.

4. Hasan F. Enzymes used in detergents: Lipases. *AFRICAN J Biotechnol*. 2010;9(31):4836 - 4844.

5. Villeneuve P, Muderhwa JM, Graille J, Haas MJ. Customizing lipases for biocatalysis: a survey of chemical, physical and molecular biological approaches. *J Mol Catal B Enzym*. 2000;9(4-6):113-148.

6. Hasan F, Shah AA, Hameed A. Industrial applications of microbial lipases. *Enzyme Microb Technol*. 2006;39(2):235-251.

7. Ghaly AE, Dave D, Brooks MS, Budge S. Production of biodiesel by enzymatic transesterification: Review: American Journal of Biochemistry and Biotechnology. *Am J Biochem Biotechnol*. 2010;6:54–76.

8. Fjerbaek L, Christensen K V, Norddahl B. A review of the current state of biodiesel production using enzymatic transesterification. *Biotechnol Bioeng*. 2009;102:1298–1315.

9. Fan X, Niehus X, Sandoval G. Lipases as Biocatalyst for Biodiesel Production. In: Sandoval G, ed. *Methods in Molecular Biology*. Totowa, NJ: Humana Press; 2012:471–483.

10. Ferrato F. A critical reevaluation of the phenomenon of interfacial activation. *LIPASES, PT B*. 1997;286:327 - 347.

11. Reis P, Holmberg K, Watzke H, Leser ME, Miller R. Lipases at interfaces: a review. *Adv Colloid Interface Sci*. 2009;147-148:237-50.

12. Mogensen JE, Sehgal P, Otzen DE. Activation, inhibition, and destabilization of Thermomyces lanuginosus lipase by detergents. *Biochemistry*. 2005;44(5):1719-30.

13. Sarda L. Action de la lipase pancreatique sur les esters en emulsion. *Biochim Biophys Acta*. 1958;30(3):513 - 521.

14. Martinelle M, Holmquist M, Hult K. On the interfacial activation of Candida-Antarctica lipase-A and lipase-B as compared with Humicola-Lanuginosa lipase. *Biochim Biophys Acta-Lipids Lipid Metab*. 1995;1258:272-276.

15. Housaindokht MR, Monhemi H. The open lid conformation of the lipase is explored in the compressed gas: New insights from molecular dynamic simulation. *J Mol Catal B Enzym*. 2013;87:135-138.

16. Stamatis H, Xenakis A, Menge U, Kolisis FN. Kinetic study of lipase catalyzed esterification reactions in water-in-oil microemulsions. *Biotechnol Bioeng*. 1993;42(8):931-7.

17. Zhou G-W, Li G-Z, Xu J, Sheng Q. Kinetic studies of lipase-catalyzed esterification in water-in-oil microemulsions and the catalytic behavior of immobilized lipase in MBGs. *Colloids Surfaces A Physicochem Eng Asp*. 2001;194:41–47.

18. Cornish-Bowden A. *Fundamentals of Enzyme Kinetics*. Portland Press; 1995.

19. Al-Zuhair S, Ling FW, Jun LS. Proposed kinetic mechanism of the production of biodiesel from palm oil using lipase. *Process Biochem*. 2007;42:951–960.

20. Verger R, Mieras MCE, Haas GH de. Action of Phospho lipase a EC-3.1.1.4 at interfaces. *J Biol Chem*. 1973;248(11):4023 - 4034.

21. Szczęsna Antczak M, Kubiak A, Antczak T, Bielecki S. Enzymatic biodiesel synthesis – Key factors affecting efficiency of the process. *Renew Energy*. 2009;34:1185–1194.

22. Brockerhoff H. Substrate specificity of pancreatic lipase. *Biochim Biophys Acta - Enzymol*. 1968;159(2):296-303.

23. Kuo TM. *Lipid Biotechnology*. Marcel Dekker; 2002.

24. Hermansyah H, Wijanarko A, Kubo M, Shibasaki-Kitakawa N, Yonemoto T. Rigorous kinetic model considering positional specificity of lipase for enzymatic stepwise hydrolysis of triolein in biphasic oil-water system. *Bioprocess Biosyst Eng*. 2010;33:787-796.

25. Pilarek M, Szewczyk KW. Kinetic model of 1,3-specific triacylglycerols alcoholysis catalyzed by lipases. *J Biotechnol*. 2007;127:736–744.

26. Roberto F-L, Fernandez-lafuente R. Lipase from Thermomyces lanuginosus: Uses and prospects as an industrial biocatalyst. *J Mol Catal B Enzym*. 2010;62:197–212.

27. Dizge N, Keskinler B. Enzymatic production of biodiesel from canola oil using immobilized lipase. *Biomass and Bioenergy*. 2008;32(12):1274-1278.

28. Chen X, Du W, Liu DH. Effect of several factors on soluble lipase-mediated biodiesel preparation in the biphasic aqueous-oil systems. *World J Microbiol Biotechnol*. 2008;24:2097-2102.

29. Al-Zuhair S, Hasan M, Ramachandran KB. Kinetics of the enzymatic hydrolysis of palm oil by lipase. *Process Biochem*. 2003;38:1155-1163.

30. Bronlund JE. Separating the Meat from the Soup. *Biotechnol Adv*. 2003;21(3):211-212.

31. Cesarini S, Diaz P, Nielsen PM. Exploring a new, soluble lipase for FAMEs production in water-containing systems using crude soybean oil as a feedstock. *Process Biochem*. 2013;48(3):484-487.

32. Kaieda M, Samukawa T, Kondo A, Fukuda H. Effect of Methanol and water contents on production of biodiesel fuel from plant oil catalyzed by various lipases in a solvent-free system. *J Biosci Bioeng*. 2001;91(1):12-15.

33. Kamal MZ, Yedavalli P, Deshmukh M V, Rao NM. Lipase in aqueous-polar organic solvents: activity, structure, and stability. *Protein Sci*. 2013;22(7):904-15.

34. Hermansyah H, Kubo M, Shibasaki-kitakawa N, Yonemoto T. Mathematical model for stepwise hydrolysis of triolein using Candida rugosa lipase in biphasic oil-water system. *Biochem Eng J*. 2006;31:125-132.

35. Al-Zuhair S, Ramachandran KB, Hasan M. Investigation of the specific interfacial area of a palm oil-water system. *J Chem Technol Biotechnol*. 2004;79:706-710.

36. Lv D, Du W, Zhang G, Liu D. Mechanism study on NS81006-mediated methanolysis of triglyceride in oil/water biphasic system for biodiesel production. *Process Biochem*. 2010;45:446–450.

37. Al-Zuhair S. Production of Biodiesel by Lipase-Catalyzed Transesterification of Vegetable Oils: A Kinetics Study. *Biotechnol Prog*. 2005;21:1442–1448.

38. Cheirsilp B, H-Kittikun A, Limkatanyu S. Impact of transesterification mechanisms on the kinetic modeling of biodiesel production by immobilized lipase. *Biochem Eng J*. 2008;42:261–269.

39. Ricca E, Gabriela M, Stefano DP, Iorio G, Calabrò V, Paola M de, Curcio S. Kinetics of enzymatic trans-esterification of glycerides for biodiesel production. *Bioprocess Biosyst Eng*. 2010;33:701–710.

40. Li W, Li R, Li Q, Du W, Liu D. Acyl migration and kinetics study of 1(3)-positional specific lipase of Rhizopus oryzae-catalyzed methanolysis of triglyceride for biodiesel production. *Process Biochem*. 2010;45:1888–1893.

41. Fedosov SN, Brask J, Pedersen AK, Nordblad M, Woodley JM, Xu X. Kinetic model of biodiesel production using immobilized lipase Candida antarctica lipase B. *J Mol Catal B Enzym*. 2013;85-86:156-168.

42. Samukawa T, Kaieda M, Matsumoto T, Ban K, Kondo A, Shimada Y, Noda H, Fukuda H. Pretreatment of immobilized Candida antarctica lipase for biodiesel fuel production from plant oil. *J Biosci Bioeng*. 2000;90:180–183.

43. Chen WW, Niepel M, Sorger PK. Classic and contemporary approaches to modeling biochemical reactions. *Genes Dev*. 2010;24(17):1861-75.

PART II
Methodology and Modelling

Chapter 4: Methodology and Modelling tools

The tools and methods used in the model based design for enzymatic biodiesel production are presented along with the methodology of the procedure

4.1. Introduction

Reaction engineering and reactor design has always been the "backbone" of the chemical engineering discipline. Detailed studies on reactor selection, heat and mass balances and mixing in large scale reactors have been discussed extensively in the scientific literature[1–3]. Heuristics for reactor selection often results in a good reactor selection in many cases. For example the use of batch data to predict reactor configuration residence times and conversions are easy to apply and fast to use.

Applying reaction engineering principles to bio-catalytic systems is not too different. However, it should be noted that there is a difference in using conventional catalysis and using a biocatalyst. Enzymes for example generally catalyse reactions at low substrate and product concentrations and are very selective[4]. Conventional catalysis generally operates at high substrate and product concentrations but is not very selective. What it then "boils down to" is that the mechanism of the reaction that is different. Hence the focus of this thesis is placed on the mechanistic modelling of the enzyme kinetics which is then coupled to the reactor mass balance to predict how the reactor should be operated. The key feature of a mechanistic model is that it is reasonable, consistent with known data and phenomena of the system. There are numerous papers in the literature that describe modelling of enzyme kinetics based on the simplest enzyme reaction mechanism consisting of a binding and a catalytic step. For kinetics with two substrates, more complex mechanistic models such as a Ping-Pong Bi-Bi reaction are postulated, but are usually based on how the reaction rate varies with substrate concentration (initial rate data). Also an issue with simple kinetics is which steps of the mechanism are kinetically significant? Estimations of the rates of elementary reactions and enzyme substrate complexes via mechanistic modelling are a consequence of the analysis and not based on steady state assumptions of the enzyme substrate complex.

However, mechanistic models tended to have numerous parameters that need to be estimated. Effective estimation of parameters in bio-catalytic kinetic expressions is very important when building process models to enable evaluation of process technology options and modes of operation. In developing the kinetic model, a work flow was followed which will be elaborated upon in this chapter.

4.2. Strategy for Development of Reaction Model

The modelling work flow is based on the work by Heitzig and co-workers[5]. The methodology is based on the concept of decomposing the modelling work into a sequence of modelling tasks and the associated methods, tools and data needed to perform such tasks. This is illustrated in Figure 4-1 where the solution of each of the steps becomes the input for the subsequent step.

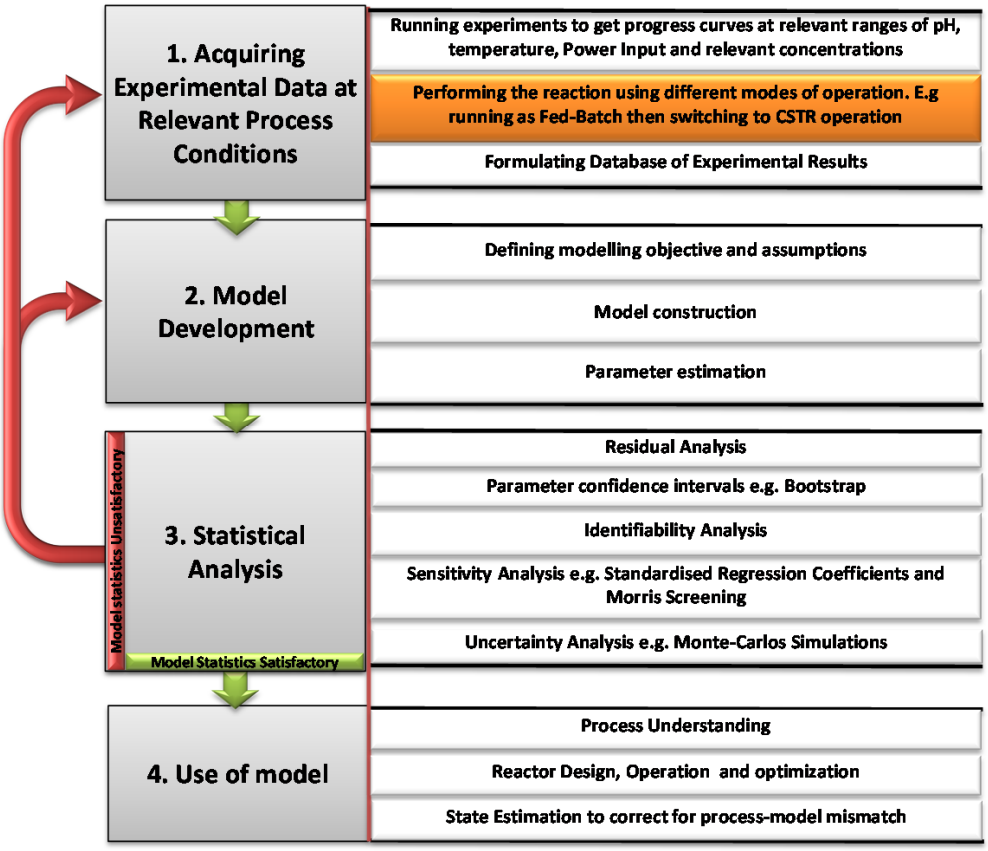


Figure 4-1 Work flow in the model development for the reaction system

4.2.1. First step: Acquiring experimental data at relevant process conditions

Collecting experimental data: For bio-catalytic process development, the enzyme reaction rate should be found under experimental conditions that apply to a potential

industrial process in question[6]. Hence, the reaction should be performed at the pH, temperature, substrate and product concentrations that the enzyme will experience. This should be done so as to evaluate the influence on the enzyme activity, rate of deactivation and hence the reaction rate over the entire course of the reaction at industrial conditions. However, in terms of mechanism development it may be necessary to perform experiments at unconventional concentrations to help elucidate a particular mechanism. For example operating at high levels of water has a negative effect on the biodiesel yield but is necessary so as to be able to discern what effects the process inputs have on the reaction.

As proposed by Al-Haque and co-workers the use of initial rate data is used in the parameter estimation step to get initial parameter estimates for a simplified form of the model; after which progress curve data are used to find the best parameter set that describes the experimental data[7]. Usually these initial rate experiments are done in the 1-2 mL scale. The results at this scale are able to show trends. However, when moving up in scale 1 to 2 orders of magnitudes the initial rates obtained are not the same, most likely due to the difference in mixing/power input. Hence it is advised to collect progress curve data at the bench or pilot scale and to increase the sampling frequency (not an issue if sampling is done online compared to offline) during the linear initial rate potion of the reaction to obtain better estimates of the initial rate.

Changing type of reactor: An interesting result obtained when investigating moving from Fed-Batch to continuous operation (see Chapter 8) was by performing the transesterification reaction using different modes of operation. Doing an experiment where the reaction was started as fed-batch then switched to CSTR operation aided in reducing the correlation of the most highly correlated parameters. This is not a method that is seen in the scientific literature to aid in the parameter estimation and is hence highlighted as a potential method developed in the thesis. Its considered potential given no theory has been developed to show that the method works on all cases or if the result was just a coincidence. Nevertheless, it is included and is definitely an area for further research.

Database of experimental data: Last but not least, in terms of experimental data collection and its use in modelling work, the use of templates to standardise the data collection amongst various collaborators and storage of the experimental data in a databases is essential. This aids in the collection and analysis of data especially when there

are many collaborators working on a project. If this is done properly data can easily be shared between groups of researchers instead of by one individual or department. The issue isn't trivial especially in the way how the activity, expression and amount of protein in an enzyme formulation may vary. However the work by Gardossi and co-workers where they lay out guidelines on how to report bio-catalytic reaction goes a long way in making these kinds of databases a reality[8].

4.2.2. Second step: Model Development

Defining modelling objectives and assumptions: In reality the first and second step are intertwined. This is so given that in the model development stage information is gathered from the scientific literature on experiments that have already been performed and the various phenomena that characterise the system (e.g. kinetic or mass transfer effects). It is during the model development step where the modelling objectives and modelling assumptions are defined. The modelling objective outlines what is the purpose of the model and what it is to be used for. The modelling assumptions help to frame/simplify the model construction step.

Model Construction: In the model construction step the conservation equations (e.g. mass and energy balance) are formulated for the system and are converted into mathematical terms. Depending on the volume balance, if it is ideally mixed, the system can mathematically be represented by algebraic equations and/or ordinary differential equations. If the volume however is distributed then partial differential equations are used. The conservation equations are then combined with the constitutive equations which are the various phenomena taking place such as reaction kinetics.

Parameter Estimation: Once the mathematical model has been formulated the unknown parameters are determined in the parameter estimation step. In this step, the cost function to be minimised can be formulated as a non-linear programming problem. All proofs, which can be found in standard textbooks on optimization are omitted from the discussion[9,10]. The mathematical models considered are assumed to be described by a system of differential algebraic equations:

Find θ to minimize

$$f = \int_0^{t_f} \left(y_{data} - y(\theta, t)\right)^T W(t) \left(y_{data} - y(\theta, t)\right) dt \quad \text{Eqn. (4.1)}$$

Subject to

$$\bar{J} = \left(\frac{dx}{dt}, x, y, \theta, v, t\right) = 0 \quad \text{Eqn. (4.2)}$$

$$x(t_0) = x_0 \quad \text{Eqn. (4.3)}$$

$$h(x, y, \theta, v) = 0 \quad \text{Eqn. (4.4)}$$

$$g(x, y, \theta, v) \leq 0 \quad \text{Eqn. (4.5)}$$

$$\theta^L \leq \theta \leq \theta^U \quad \text{Eqn. (4.6)}$$

Where

f - Cost function to be minimized

θ - Vector of unknown parameters

y_{data} - Experimental data at a time t

$y(\theta, t)$ - Model prediction at the time t

$W(t)$- Weighting or scaling matrix

x - Differential state variables

v - Vector of other (usually time-invariant) parameters that are not estimated

$\bar{J}$ - Set of differential and algebraic equality constraints describing the nonlinear process model

h and g - Possible equality and inequality constraints

From the problem formulation an optimization algorithm is needed to minimize f; and a differential equation solver is needed to solve the underlying set of differential equations. An efficient solution of the differential equations is hence crucial to the performance of the overall algorithm. It should be noted there is no universal optimization algorithm. Rather, there are numerous algorithms, each of which is tailored to a particular type of optimization problem. It is often the user's responsibility to choose an algorithm that is appropriate for their specific application. This choice is an important one; it may determine whether the problem is solved rapidly or slowly and, indeed, whether the solution is found at all[9]. The same goes for the solution of the differential equations in the model.

The integration method chosen for a specific application should reflect the nature of the problem at hand. For example, Runge-Kutta methods are an important family of implicit and explicit iterative methods which are most commonly used. If the problem is stiff, an implicit integrator with strong stability properties should be used. If it is not stiff, the use of an explicit method is adequate. When frequent discontinuities are present, one-step methods should be used, whereas multi-step methods are advantageous for problems with long and smooth intervals. If high accuracy in the solution is required, a method with high order should be chosen. These guidelines, however, are not always useful, since for example the stiffness characteristics of a specific problem are often not known beforehand. Consequently, for reasons of reliability and robustness, implicit methods are often the default choice for many practical purposes.

4.2.3. Third step: Statistical Analysis

This step is where the model is analysed to determine if the model could be used in step four for what it was designed for or if it is we need to return to step one or two for further development of the model. This working cycle is iterated until a termination criterion, e.g. on the accuracy, is fulfilled

Residual Analysis: Granted there are no numerical issues, the unknown parameter values obtained are determined to sufficient precision based on the experimental data, in the parameter estimation step. One quick way to visually inspect how well the model fits the data is by analysis of the residuals. For the model to describe the system sufficiently, the residuals should behave as Gaussian white noise with zero mean and a finite variance. In this work the histogram of residuals is used to gauge if the model complexity is sufficient or if during the parameter estimation step the data is being over-fitted in which case the histogram is usually skewed.

Confidence intervals: In general when parameter estimates are being reported, the confidence intervals should also be added. This will aid in being able to determining the reliability of the parameter estimates obtained along with being able to be able to do further analysis such as uncertainty and sensitivity analysis. Given the data has measurement errors which are usually assumed to be normally distributed, with a known variance, it is then possible to estimate the variance of the parameters using a bootstrap method[11]. A bootstrap method is more robust method to calculate realistic confidence

intervals for the parameter estimates for dynamical systems compared to methods based on maximum likelihood estimation and is hence used in this work[12,13].

The Bootstrap Method uses the residuals randomly picked from the least squares fit to generate synthetic datasets, which are then fit using the same least squares algorithm as used on the actual data. N number of datasets are synthesised and refitted, giving N sets of parameters. By the Central Limit Theorem, it is assumed the sample mean of the bootstrapped parameter estimates are normally distributed. From these sets the confidence intervals for the parameters can be estimated.

Identifiability analysis: The goal of performing an identifiability analysis is to identify parameter subsets that are noncollinear and therefore identifiable. The identifiability method based on the work by Brun and co-workers was used to ascertain which parameters could actually be identified from the available experimental data given the model structure[14].

Uncertainty and Sensitivity analysis: Given some of the parameters in the model may not be identifiable it is then imperative to ascertain how the uncertainty in the parameter estimated affects the model outputs. For the uncertainty analysis, the standard Monte Carlo procedure is used to propagate and analyse the uncertainty in the model parameters. This method offers global results due to the large number of model evaluations performed using the randomly sampled parameters, to obtain the distribution of the model outputs.

The purpose of the sensitivity analysis is to identify parameters and variables that have no impact on the desired model output. This information might be used for model simplification or to go back and design new experiments to be able to identify the non-sensitive parameters. Different methods for sensitivity analysis exist. To evaluate and rank the output variance of the model with respect to the model parameters, Standard Regression Coefficients and Morris screening are used for the sensitivity analysis. Two sensitivity analysis methods are used, to make it possible to corroborate the results obtained. If both methods identify the same parameters then it gives greater confidence in the results obtained.

4.2.4. Fourth step: Use of the Model

The whole procedure in the mechanistic model development gives great insight into how the system behaves given one needs to “dig into” the phenomena occurring in the system. Once an acceptable model is developed, the model is then used to guide how the reactors should be operated. Since experiments are often expensive and time-consuming, the insight gleaned from the model is valuable. Also, the data obtained from the additional experiments can be used to improve the estimates of the model parameters. In fact, the model can be used to devise a set of experiments that yield parameter estimation with maximum statistical quality, e.g. smallest possible confidence intervals for the parameters.

All too often however the model may not have the accuracy that is desired. In Chapter 9 the use of a Continuous-Discrete Extended Kalman Filter (a state estimator) is used to correct for mismatch between the process data and the process model. What we demonstrate is that with our imperfect model, coupled to measurements of the system in the Continuous-Discrete Extended Kalman Filter framework, we can get corrected estimates of our states. The filter is relatively easy to tune given the single tuning parameter. This then lays the foundation for use of the model in a model based control framework given that it is possible to get accurate predictions of our components in the reactor, for various changes to the process inputs. The state estimator can then be used to identify outliers and help filter the measurement data. The ability to correct for the process-model mismatch and identify outliers in the measurement data will prove useful in any process monitoring framework.

4.3. Conclusions

Mathematical models have some significant limitations which need to be kept in mind during the modelling cycle:

- Quantity and accuracy of the available data. The success and results of mathematical models depend largely on the experimental data. A mathematical model cannot be better than the physical or chemical data on which it is based.
- The accuracy required for the individual parameters of the model depends on the sensitivity with which the results of the model respond to changes in these parameters. Those parameters which exert the greatest influence on the results of the model must be determined with the greatest accuracy.

- Complex reaction mechanisms can lead to many model parameters. Consequently, the modelling effort can be difficult and time consuming.
- The model structure itself may also be inappropriate. This can result from incorrect assumptions or erroneous simplification of the model.
- Complex structures can often be formulated mathematically which are usually non-linear. Estimating parameters in such systems is both computationally intensive as well as numerically challenging due to a variety of undesirable characteristics, such as poor initial guess for parameters, ill-conditioning and stiffness of model equations.

However, the advantages in mathematical modelling far outweigh the disadvantages and are a key technique to gain insight into the dynamics of a system. The main points are:

- Existing processes can be investigated more quickly and economically
- Sensitivity analysis can be investigated by varying variables and parameters, negating the need for the repetition of tests/experiments.
- Optimization of complex systems can be determined in accordance with changing requirements and regulations
- Extrapolation to test extreme operating conditions that are not possible or practical for the physical experiment. Extrapolation of data must, however, always be discussed in the light of the limits of the model

4.4. References

1. Levenspiel O. *Chemical Reaction Engineering*. Wiley; 1999.
2. Perry RH. *Perry's Chemical Engineers' Handbook*. McGraw-Hill; 2007.
3. Fogler HS. *Elements of Chemical Reaction Engineering*. Prentice Hall PTR; 2005.
4. Schmid A, Dordick JS, Hauer B, Kiener A, Wubbolts M, Witholt B. Industrial biocatalysis today and tomorrow. *Nature*. 2001;409(6817):258-68.
5. Heitzig M, Sin G, Sales-Cruz M, Glarborg P, Gani R. Computer-Aided Modeling Framework for Efficient Model Development, Analysis, and Identification: Combustion and Reactor Modeling. *Ind Eng Chem Res*. 2011;50:5253–5265.
6. Vasic-Racki D, Kragl U, Liese A. Benefits of Enzyme Kinetics Modelling *. *Chem Biochem Eng Q*. 2003;17(1):7-18.

7. Al-Haque N, Santacoloma PA, Neto W, Tufvesson P, Gani R, Woodley JM. A robust methodology for kinetic model parameter estimation for biocatalytic reactions. *Biotechnol Prog*. 2012;28:1186–1196.

8. Gardossi L, Poulsen PB, Ballesteros A, Hult K, Švedas VK, Vasić-Rački Đ, Carrea G, Magnusson A, Schmid A, Wohlgemuth R, Halling PJ. Guidelines for reporting of biocatalytic reactions. *Trends Biotechnol*. 2010;28:171–180.

9. Nocedal J, Wright SJ. *Numerical Optimization*. 2nd ed. New York: Springer; 2006:1 online resource (xxii, 664.

10. Rao SS. *Engineering Optimization*. Hoboken, NJ, USA: John Wiley & Sons, Inc.; 2009:1 - 813.

11. Efron B. 1977 Rietz lecture - Bootstrap methods - Another look at the jackknife. *Ann Stat*. 1979;7:1-26.

12. Rodriguez-Fernandez M, Banga JR, Doyle FJ. Novel global sensitivity analysis methodology accounting for the crucial role of the distribution of input parameters: application to systems biology models. *Int J Robust Nonlinear Control*. 2012;22:1082–1102.

13. Dogan G. Bootstrapping for confidence interval estimation and hypothesis testing for parameters of system dynamics models. *Syst Dyn Rev*. 2007;23(4):415-436.

14. Brun R, Kuhni M, Siegrist H, Gujer W, Reichert P. Practical identifiability of ASM2d parameters - systematic selection and tuning of parameter subsets. *Water Res*. 2002;36:4113-4127.

Chapter 5: Uncertainty and Sensitivity Analysis Applied to Enzymatic Biodiesel Production

In this chapter the application of uncertainty and sensitivity analysis is applied to a kinetic model given the inherent uncertainty in the kinetic parameters.

A modified version of this chapter where these methods were applied to another kinetic model has been accepted for publication in the Proceedings of 12th IFAC Symposium on Computer Applications in Biotechnology as Price, J. A., Nordblad, M., Woodley, J., & Huusom, J. K. (2013). Application of Uncertainty and Sensitivity Analysis to a Kinetic Model for Enzymatic Biodiesel Production

5.1. Introduction

The process conditions and the model parameters (e.g. kinetic parameters) are the main components that influence the prediction quality of a model[1]. In a host of various engineering fields, the standard Monte Carlo procedure has been used to statistically analyse the effect of uncertainty in the input factors (model parameters and/or process conditions) on the model outputs (uncertainty analysis); along with sensitivity analysis based on variance decomposition to identify and quantify which input factors were most influential to the model outputs[2–5].

One shortfall with the existing kinetic models for enzymatic biodiesel production presented in Chapter 3, is that to the best of our knowledge, none that describe the enzymatic transesterification have been statistically analysed to ascertain the working bounds of the model. This remains as a weak point in the credibility of these kinetic models and hence their applicability for engineering design. The aforementioned points set the stage for this study.

In this chapter the kinetic model presented by Cheirsilp and co-workers, is used as a case study[6]. The aim is to:

- Characterize the regions where the model is most reliable under uncertainty in the parameter estimates
- Identify which parameters contribute most to the uncertainty in the model outputs
- Simply the kinetic model if there are insensitive parameters

The Uncertainty Analysis (Monte Carlo simulations) enables for a predetermined parameter uncertainty the evaluation of the working bounds of the model. The outcome is a better understanding of the predictive accuracy of the model during the course of reaction. The Sensitivity Analysis (Standard Regression Coefficients and Morris screening) enables the identification of the group of influential and non-influential parameters to the model outputs. The influential parameters will help identify which parameters contribute most to the variance in the predicted concentrations of the model outputs. From this it can be deduced what mechanisms dominate at a particular point in the reaction. The non-influential parameters have negligible contribution to the variance in the model outputs.

Hence these parameters can be fixed within the range of the parameter variability, aiding in model simplification.

This chapter is organised as follows. The methodology used is introduced, followed by the theory for the uncertainty/sensitivity analysis techniques used. The case study is presented, along with the methods used in the simulations. The results from the uncertainty analysis are discussed, followed by the discussion of the results from the sensitivity analysis of the Standard Regression Coefficients and Morris screening. The results of the uncertainty and sensitivity analysis are then put into perspective for use in further engineering work.

5.2. Methodology

To help improve modelling processes Foss and co-workers investigated the process of model development in chemical industries and laid out some guidelines to improve modelling technology[7]. More recently Heitzig and co-workers proposed a generic methodology that structures the process of model development and analysis[8]. This coupled with the work done by Sin and co-workers where statistical tools are used during the modelling process are combined in the methodology used in this work[2]. The proposed methodology is illustrated in Figure 5-1.

Defining the modelling objectives is essential in framing the goals and expected outcomes from the model. The information gathering process entails the collection of relevant experimental data along with phenomena occurring in the system such as reaction kinetics. To aid in this step, advice from experts, experimental observations and literature reviews are essential. During the model construction phase, modelling assumptions were made to help frame the problem. Mass and energy balances are made around the system boundaries and the constitutive equations relating to the underling process phenomena are mathematically formulated. For the uncertainty analysis, the standard Monte Carlo procedure is used to propagate and analyse the uncertainty in the model parameters. To evaluate and rank the output variance of the model with respect to the model parameters, Standard Regression Coefficients and Morris screening are used for the sensitivity analysis. Two sensitivity analysis methods are used, to make it possible to corroborate the results obtained. If both methods identify the same parameters then it gives greater confidence in the results obtained.

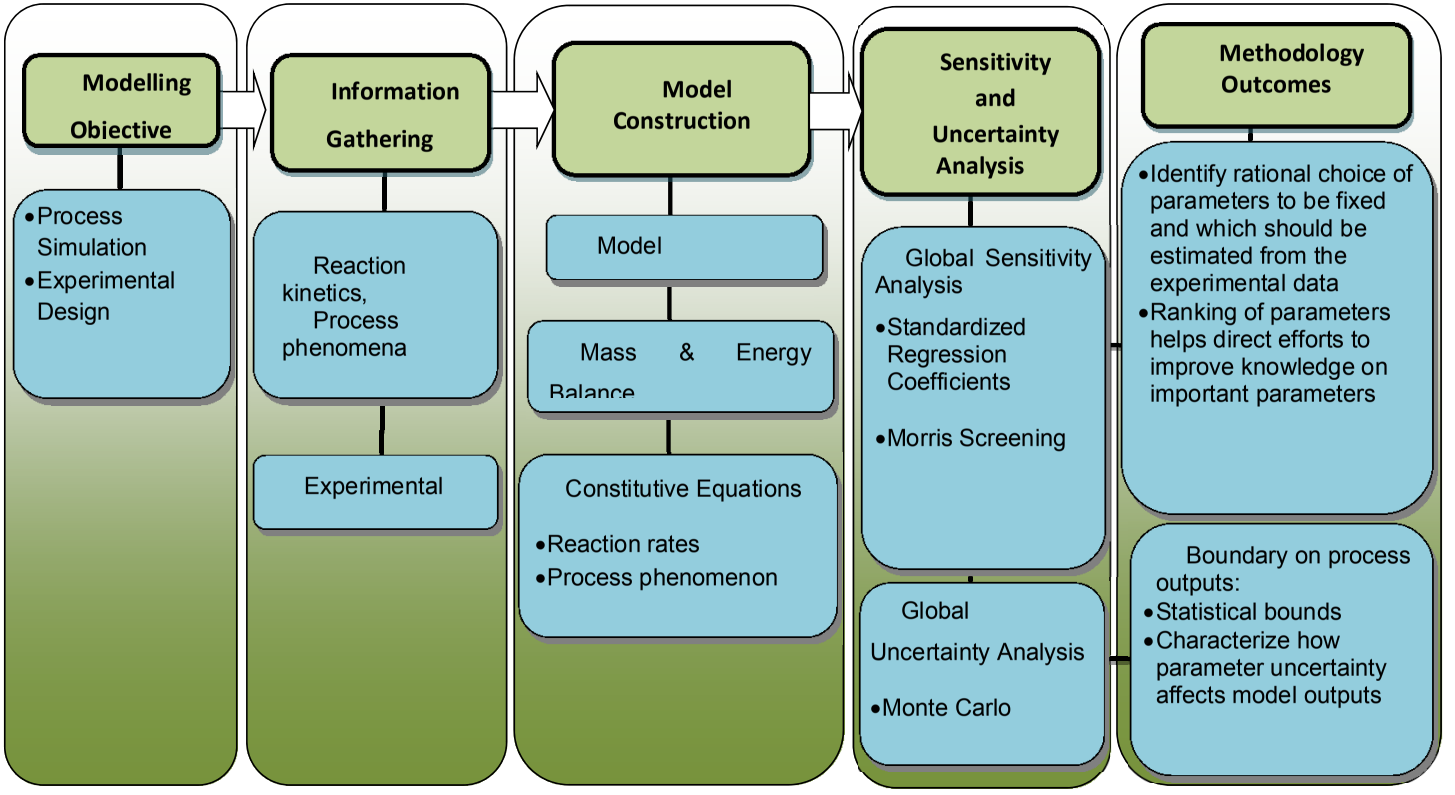


Figure 5-1 Methodology used in this work for the Sensitivity and Uncertainty analysis

5.3. Theory

In the following section, a description of the main tenants of each method used for the uncertainty and sensitivity analysis is highlighted. All proofs can be found in standard textbooks on parameter estimation and statistics and have been omitted[9–11].

5.3.1. Uncertainty analysis: Monte-Carlo simulations

Uncertainty analysis aids in the decision making process, by helping to put bounds on the model outputs, for a specified uncertainty in the model factors. In this work, the Monte Carlo technique was used in the evaluation of the kinetic parameters uncertainty, on the model outputs. The method offers global results due to the large number of model evaluations that are required in the analysis. The method samples from the input parameter space, θ and generate model outputs, *y*. The Monte Carlo analysis of uncertainty involved three steps:

1. Specifying input uncertainty.
2. Sampling input uncertainty.
3. Simulating the model using the sampling matrix to obtain prediction uncertainty for *y*.

Specifying input uncertainty: Specifying input uncertainty for the parameters is no trivial task. If a parameter estimation procedure was performed, the confidence intervals of the parameters could be calculated and used as the upper and lower bound of a kinetic parameter. However, for the case where the model with parameter estimates $\hat{\theta}$ is provided but parameter estimation is not carried out due to insufficient experimental data; it is then necessary to use expert review and/or consult the relevant literature resources about the uncertainty of the parameter estimates.

Sampling input uncertainty: Latin hypercube sampling with correlation control has become a widely used sampling technique for the propagation of uncertainty in analyses of complex systems. This is due to its ease of implementation and dense stratification over the range of each sampled variable while maintaining the correlation structure[12]. Samples are selected from the input parameter space, where each sample, θ_i contains one value for each input parameter creating a $\left[M \times N_{lh}\right]$ matrix. Where *M* stands for the total number of model parameters, and N_{lh} is the total number of Latin-Hypercube samples.

Simulating the model using the sampling matrix: N_{lh} dynamic simulations were then performed using the $\left[M \times N_{lh}\right]$ sampled input matrix. Each simulation result is then stored in a $\left[N_t \times N_u \times N_{lh}\right]$ size array where, N_t is the length of the discrete time series and N_u is the number of model outputs. The complete Monte Carlo results provide a cumulative distribution function for each output variable at each time instant. The uncertainty of the model outputs is then represented using mean and percentile calculations.

5.3.2. Sensitivity Analysis techniques – SRC and Morris screening

The purpose of sensitivity analysis is manifold, and is complimentary to uncertainty analysis[1]. Sensitivity analysis can be used to:

- Analyze the output variance of a model with respect to input factors.
- Rank the most significant parameters affecting the model outputs.
- As an effective tool for mechanism identification and model reduction.

The sensitivity analysis methods used are Standardised Regression Coefficients along with Morris screening.

Standardised Regression Coefficients, linear regression of Monte-Carlo simulations: In this method the sensitivity measure was obtained by performing a linear regression for each of the model outputs of interest obtained from the Monte Carlo procedure. The method requires a scalar output, which can characterize the dynamics of the model output. For example, all biodiesel concentrations at a fixed time point can be characterized by its mean, y_m. If this is done for all model outputs, y_m will be a $\left[N_t \times N_u\right]$ sized matrix. A linear regression model is then built for each model output N_u:

$$y_{m_{i,k}} = b_{0,k} + \sum_{j=1}^{M} b_{j,k}\theta_{i,j} + \varepsilon_{i,k} \text{ for} \begin{bmatrix} i = 1, 2, ..., N_t \\ k = 1, 2, ..., N_u \end{bmatrix} \quad \text{Eqn. (5.1)}$$

Where $b_{j,k}$ is the coefficient of the j^{th} model parameter for the k^{th} model output, $\theta_{i,j}$ is the value of the j^{th} model parameter and $\varepsilon_{i,k}$ is the error of the regression model. The above equation can also be written in a dimensionless form by scaling the outputs and the parameters using their corresponding mean, μ and standard deviation, σ:

$$\frac{y_{m_{i,k}} - \mu_{(y_{m_k})}}{\sigma_{(y_{m_k})}} = \sum_{j=1}^{M} \beta_{j,k} \frac{\theta_{i,j} - \mu_{x_j}}{\sigma_{x_j}} + \varepsilon_{i,k} \qquad \text{Eqn. (5.2)}$$

$\beta_{j,k}$ is called the standardized regression coefficient of the j^{th} model parameter, θ_j for the k^{th} model output of y_{m_k}. The standardized regression coefficient is then found by solving the linear system of equations.

The advantage of using the Standard Regression Coefficients is that the effectiveness of the regression model can be immediately verified by the model coefficient of determination, R^2. For the Standard Regression Coefficients to be considered a valid measure of sensitivity, R^2 should be greater than 0.7. The sensitivity measure, β_{j,N_u} has the following characteristics[1,12]:

- β_{j,N_u} can take on values between -1 and +1.
- A high absolute value indicates a large effect of the corresponding parameter on the output.
- The sign of the coefficient indicates the effect of the parameter on the output e.g. a positive sign indicates a positive effect on the output.
- Coefficients close to zero signify that the parameter has a negligible effect on the output.

Morris Screening: This method is a one-factor-at-a-time method, meaning that in each run only one input factor is given a new value. It facilitates a global sensitivity analysis by making a number of local changes at different points. It relies on estimating the distribution of the Elementary Effects of each input factor on the k^{th} model output called $EE_{j,k}$. The method gives a good compromise between accuracy and efficiency[11]. Calculation of one elementary effect for each input requires *(M+1)*model simulations. Given r, repetitions are needed (typically 10–50), the total number of model simulations needed becomes $r \times (M+1)$. Each parameter, θ_j can only take values corresponding to p levels from its range (imagine a grid in which the range of each parameter is subdivided into p levels). The value for p could be 4, 6, and 8 which corresponds to the 25^{th}, 17^{th}, and 12.5^{th} percentile of the uniform distribution of the input factors. This distribution function is denoted as $F_{j,k}$, which stands for the distribution of the effects of the j^{th} input factor on the k^{th} output. Because this method also requires scalar values, the scalar model output matrix, y_m was used as data

for the Morris screening. The $EE_{j,k}$ attributable to each input factor is obtained from the following differentiation of model output, y_{m_k} with respect to the model parameter, θ_j:

$$EE_{j,k} = \frac{\partial y_{m_k}}{\partial \theta_j} = \frac{y_{m_k}\left(\theta_1,\ \theta_2,\ \theta_j + \Delta, \ldots,\ \theta_M\right) - y_{m_k}\left(\theta_1,\ \theta_2,\ \theta_j, \ldots,\ \theta_M\right)}{\Delta} \quad \text{Eqn. (5.3)}$$

Where Δ is a predetermined perturbation factor of θ_j, $y_{m_k}\left(\theta_1,\ \theta_2,\ \theta_j, \ldots,\ \theta_M\right)$ is the scalar model output evaluated at $\left(\theta_1,\ \theta_2,\ \theta_j, \ldots,\ \theta_M\right)$, whereas $y_{m_k}\left(\theta_1,\ \theta_2,\ \theta_j + \Delta, \ldots,\ \theta_M\right)$ is the scalar model output corresponding to a Δ change in θ_j. The choice of perturbation factor, Δ is optimal when $\Delta = p/\ 2\left(p-1\right)$.

The scaled elementary effects (using the standard deviation for the inputs, σ_{θ_j} and outputs, $\sigma_{y_{m_k}}$ respectively) in *Eq. (5.4)* is then used for comparing and ranking of the parameters.

$$\left(EE_{j,k}\right)_{scaled} = EE_{j,k}\frac{\sigma_{\theta_j}}{\sigma_{y_{m_k}}} \quad \text{Eqn. (5.4)}$$

5.4. Case study: Kinetic modelling of enzymatic biodiesel production

Various kinetic models for enzymatic transesterification of vegetable oils have been proposed[6,13–17]. The kinetic model by Cheirsilp and co-workers however, is interesting to work with given the characteristics of the kinetic model[6]. From their reported results, the

Table 5-1 Potential uses of the kinetic model by Cheirsilp and co-workers given the model characteristics.

Kinetic Model Characteristics	Modelling Outcomes
Kinetics describing how the reactants and products of interest (TAG, DAG, MAG, FFA, FAAE, Water, Glycerol, and Alcohol) vary during the entire reaction.	For a given alcohol/oil molar ratio determine when the reaction is complete
	Estimate how the changes in water and FFA concentrations affect the course of the reaction
Kinetic model includes the enzyme concentration in its mathematical expression	Estimate the required enzyme concentration to achieve a desired biodiesel yield in a specific time

characteristics of the kinetic model (Table 5-1) provide specific outcomes that are necessary to aid in process development. Hence their kinetic model is used to illustrate the application of the methodology.

5.4.1. Kinetic model overview

The kinetic model by Cheirsilp and co-workers describes the transesterification reaction for an immobilized lipase on a micro-porous polypropylene support, Lipase PS (from Pseudomonas sp)[6]. Their third proposed mechanism (see Figure 5-2) is used for the case study given it represented the experimental data the best. The mechanism is divided into two parts; the hydrolysis step to produce FFA prior to the esterification step and the ethanolysis reaction to directly produce FAAE, with the two steps occurring in parallel. The model also contains competitive alcohol inhibition.

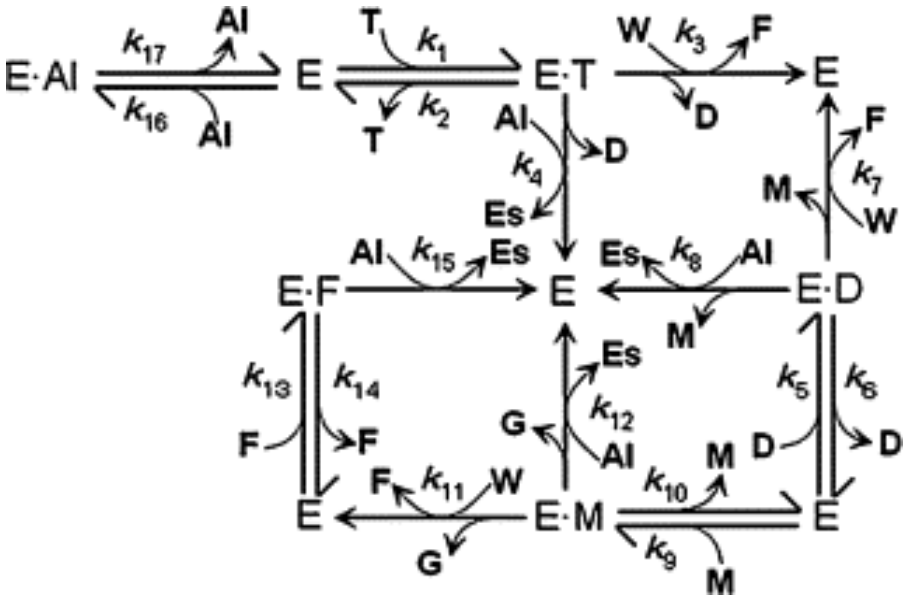


Figure 5-2 Conceptual overall reaction mechanism[6].

The kinetic parameters are taken from the authors work and further description of the model can be found in[6]. For completeness, the mathematical formulation of the model used, along with the kinetic parameters are presented in Table 5-2 and

Table **5-3**. It should be noted that Cheirsilp and co-workers base their concentrations on a mass basis. Hence all concentrations reported will also be on a mass basis.

Table 5-2 Differential equations for the batch transesterification reaction[6].

Component	Differential equations
Triglyceride [T]	$\frac{d[T]}{dt} = -\left(V_{mT}[W]+V_{eT}[Al]\right)[T][E^*]$
Diglyceride [D]	$\frac{d[D]}{dt} = \left(\left(V_{mT}[W]+V_{eT}[Al]\right)[T]-\left(V_{mD}[W]+V_{eD}[Al]\right)[D]\right)[E^*]$
Monoglyceride [M]	$\frac{d[M]}{dt} = \left(\left(V_{mD}[W]+V_{eD}[Al]\right)[D]-\left(V_{mM}[W]+V_{eM}[Al]\right)[M]\right)[E^*]$
Glycerol [G]	$\frac{d[G]}{dt} = \left(V_{mM}[W]+V_{eM}[Al]\right)[M][E^*]$
Free Fatty Acid [F]	$\frac{d[F]}{dt} = \left(\left(V_{mT}[T]+V_{mD}[D]+V_{mM}[M]\right)[W]-\left(V_{eEs}[F][Al]\right)\right)[E^*]$
Biodiesel [B]	$\frac{d[B]}{dt} = \left(V_{eT}[T]+V_{eD}[D]+V_{eM}[M]+V_{eEs}[F]\right)[Al][E^*]$
Water [W]	$\frac{d[W]}{dt} = -\left(V_{mT}[T]+V_{mD}[D]+V_{mM}[M]\right)[W][E^*]$
Alcohol [Al]	$\frac{d[Al]}{dt} = -\left(V_{eT}[T]+V_{eD}[D]+V_{eM}[M]+V_{eEs}[F]\right)[Al][E^*]$
Free Enzyme	$[E^*] = \frac{[E_T]}{\left(1+K_{mT}[T]+K_{mD}[D]+K_{mM}[M]+K_{mF}[F]+\left(\frac{[Al]}{K_i}\right)\right)}$
	Model outputs: 8
	Kinetic Parameters:12
Where $V_{mT}=\frac{k_3k_1}{k_2}, V_{mD}=\frac{k_7k_5}{k_6}, V_{mM}=\frac{k_{11}k_9}{k_{10}}, V_{eEs}=\frac{k_{15}k_{13}}{k_{14}}, V_{eT}=\frac{k_4k_1}{k_2}, V_{eD}=\frac{k_8k_5}{k_6}, V_{eM}=\frac{k_{12}k_9}{k_{10}},$ $K_{mT}=\frac{k_1}{k_2}, K_{mD}=\frac{k_5}{k_6}, K_{mM}=\frac{k_9}{k_{10}}, K_{mF}=\frac{k_{13}}{k_{14}}, K_I=\frac{k_{17}}{k_{16}}$	

Table 5-3 Parameter values used in the simulation [6]

Factor	$\hat{\theta}$
Rate Constants [mmol^{-1}h^{-1}]	
V_{mT} – Hydrolysis TAG	7.619×10^{-2}
V_{mD} – Hydrolysis DAG	8.128×10^{-2}
V_{mM} – Hydrolysis MAG	1.951×10^{-1}
V_{eT} – Ethanolysis TAG	2.751
V_{eD} – Ethanolysis DAG	1.176
V_{eM} – Ethanolysis MAG	0.965
V_{eEs} – Esterification FFA	1.383
Equilibrium Constants [g mmol^{-1}]	
K_{mT} – Equilibrium constant TAG	2.891×10^{-2}
K_{mD} – Equilibrium constant DAG	2.322×10^{-2}
K_{mM} – Equilibrium constant MAG	1.974×10^{-2}
K_{mF} – Equilibrium constant FFA	1.121×10^{-2}
Ethanol inhibition Constant [mmol g^{-1}]	
K_i	0.882

Model assumptions. The model was based on the following general assumptions:

1. Rapid equilibrium of the enzyme substrate complexes.
2. Irreversible esterification and ethanolysis reactions.
3. Perfect mixing. The whole reaction system could be regarded as a quasi-homogeneous system (Lumped model).
4. Reaction is rate controlled and mass transfer effects are ignored[15].
5. No enzyme deactivation.

The assumptions made by Cheirsilp and co-workers in their model formulation, helps in simplifying the reaction system, while capturing the main phenomenon seen in their system[6]. However it must be noted that this limits what the model can be used for. A review of enzymatic biodiesel production by Fjerbaek and co-workers show many instances where the biodiesel yield is below 90 % at the end of the reaction[18]. For a case where the equilibrium yield at the end of the reaction is not stoichiometric, the model will not be able to be extended to such a case due to the assumption made in 2). The assumption made in 4) ignores the difference between the interfacial and bulk concentrations of the enzyme, substrates and products at the oil-water interface. This means any parameters found are for a particular mixing regime, more specifically, the oil-water interfacial area produced from the mixing. It is also known that irreversible enzyme deactivation can occur depending on the alcohol concentration and repeated use[14]. The assumption made in 5) limits the use of

Table 5-4 Simulation settings used for the uncertainty and sensitivity analysis

Uncertainty analysis	Sensitivity Analysis	
Monte Carlo simulations	**Standardised Regression Coefficients**	**Morris Screening**
Sampling input uncertainty 500 samples were selected. Parameters were considered to be uncorrelated due to unavailability of the information on the correlation matrix.	The R^2 value for the linear regression for each of the model outputs of the Monte Carlo simulations should be greater than 0.7	The number of levels, p and number of repetitions , r were defined as 6 and 30, respectively
Input uncertainty of ± 50 % variability around the parameter estimates is used.	The mean of the model outputs from the Monte Carlo simulations are used at a time of 5 hrs. to enable comparison of the ranking of the parameters for the two sensitivity methods investigated	
All model parameters were assumed to have a uniform probability distribution		

the model in the evaluation of multiple batch reactions, where it is known repeated use of the enzyme shows a decrease in enzyme activity.

5.5. Uncertainty and sensitivity analysis simulation settings

The settings used for the uncertainty and sensitivity analysis, are shown in Table 4. It should be noted that the approach by Sin and co-workers in classifying the input factor uncertainty, is to class the parameters into three different groups, depending on their level of uncertainty[2,3]. Given there are no other reported parameters for this reaction, the third class is chosen, which has the highest uncertainty of 50% variability around the parameter estimates. This is an initial assumption that allows evaluation of how the uncertainty in the parameters estimates influences the model outputs. Also given Cheirsilp and co-workers performed a local sensitivity analysis where they used ± 50 % variability in their parameter estimates a comparison between the local and global sensitivity analysis can be made[6].

5.6. Results and Discussion

5.6.1. Monte Carlo Simulations

The uncertainty in the model outputs is represented using the mean along with the 5th and 95th percentile of the distribution of each model output, obtained from the dynamic simulation of the 500 Latin hypercube samples. For the kinetic model investigated, only the typically measured variables (TAG, DAG, MAG, FFA, FAAE) during the transesterification reaction are reported.

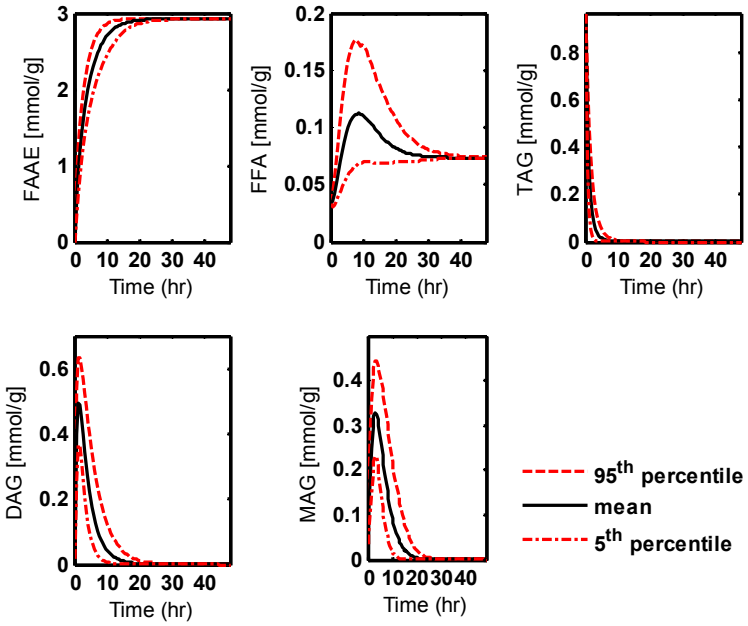


Figure 5-3 Uncertainty analysis of the model predictions for FAAE, FFA, TAG, DAG, MAG (The mean and the 5th and 95th percentiles are obtained from performing 500 Monte Carlo simulations)

Monte Carlo results: The Monte Carlo method was used to propagate the uncertainty of the kinetic parameters on the output uncertainty of the model. The interpretation of these results in Figure 5-3 is straightforward; the wider the uncertainty band (95th and 5th percentiles), the greater the influence of the parameters on the model outputs.

The magnitude of the model uncertainty differed depending on the model output. For example, the uncertainty on FAAE and TAG was relatively smaller compared to the uncertainty on the predictions of FFA, DAG and MAG. Furthermore, the uncertainty was observed to be changing over time during the reaction. The uncertainty analysis gives insight into the model structure. For all model outputs, the output uncertainty starts off small, grows and then shrinks. These phenomena could be explained by the fact that as the reaction proceeds the concentrations of the other components become more pronounced and the uncertainty of the model parameters affects the model outputs more. At the end of the reaction, the concentrations decrease, so the contribution from the parameters also decrease and hence cause a decrease in the uncertainty of the model outputs. This concept is reinforced if the TAG concentration profile is investigated. Given the TAG model

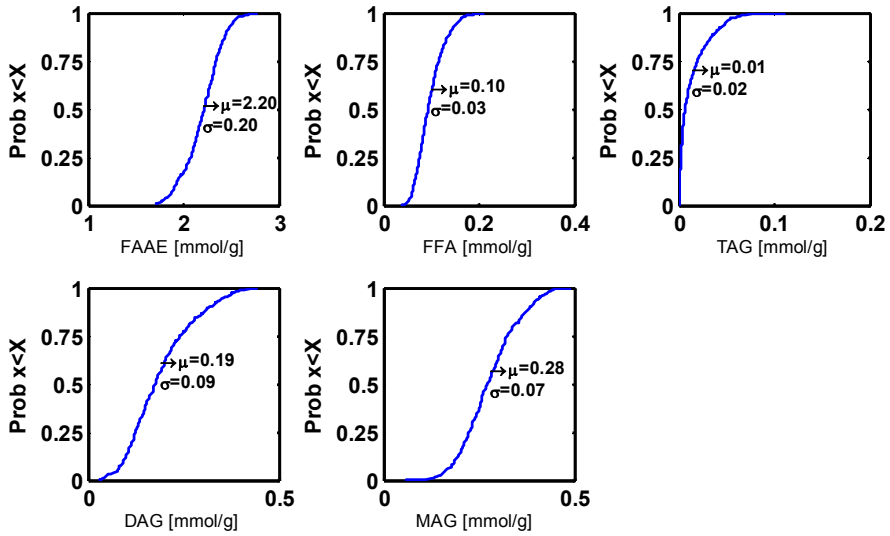


Figure 5-4 Cumulative distribution function of the 5 model outputs at a time of 5 hrs. The x-axis plots show the concentration of the model outputs and the y-axis shows the probability of a value in the x-axis being greater or equal to a chosen concentration. The mean along with the standard deviation for each model output is shown.

formulation has the fewest parameters; it also has the smallest deviation throughout the entire course of the reaction. For the ± 50 % in the parameter estimates, Figure 5-3 gives a clear depiction of where the model is most sensitive to variations in the model parameters; between the 4 hour and 25 hour part of the reaction. It should be noted, as can be seen in Figure 5-3, the reaction goes to completion. This is due to the model assumption of the esterification and ethanolysis reactions being irreversible and hence limits the application of the model to other systems.

The Monte Carlo simulation in Figure 5-3 gives the modeller much more insight into how the model behaves compared to using a local sensitivity analysis method. Statistically speaking, the uncertainty bands observed in Figure 5-3, correspond to the distribution of the model outputs at each time instant. A look at the Cumulative distribution function, paints a better picture of the acceptability of the model outputs. In Figure 5-4, the variance in the model outputs changes over the course of the reaction. Any time point can be chosen but for illustrative purposes a time period of 5 hours is used. The Cumulative distribution function in Figure 5-4 shows that the DAG concentration has a mean value of 0.19 mmol/g with a standard deviation of 0.09 mmol/g. This is a quite wide variation compared to the

FAAE concentration which has a mean value of 2.20 mmol/g with a standard deviation of 0.20 mmol/g. Depending on the application this may or may not be acceptable.

Given the ± 50 % variability used on the parameter estimates, the decision maker now has statistical meaningful bounds on which to base further calculation. Take for example, an engineer, who wishes to do an economic evaluation on the final FAAE yield. At the reaction end time of 48 hrs the FAAE concentration has a mean value of 2.93 mmol/g with a standard deviation of 0.01 mmol/g.

5.6.2. Sensitivity Analysis

For the calculation of the Standardised Regression Coefficients and the Morris screening a scalar output was needed. Given our interest is in determining which parameters can be attributed to influencing the large variability in the model outputs, a time period where there are significant variations in the model outputs of the Monte Carlo simulations (Figure 5-3) is chosen (between 5 - 15 hours). The time of 5 hours is chosen for the analysis. This is done so as to compare the rankings of the parameters obtained, from the two sensitivity analysis methods. It should be noted the analysis can be performed at different time points in which case the parameter ranking can vary.

Standard Regression Coefficients results: The degree of linearization indicated by the coefficient of model determination (R^2 value) obtained from the linear least squares fitting of *Eq. (5.2)* was over 0.8 (The detailed values can be seen in Table 5-5) This indicates that the linearized model was able to explain most of the variance in the five model outputs investigated, and hence, the corresponding coefficients can reliably be used to assess the importance of the kinetic parameters on the model outputs.

The Standard Regression Coefficients were ranked for each output, and a summary of the ranking is given in Figure 5-5. Analysing the model outputs shows that the FAAE model output is most influenced by the alcohol inhibition K_i along with V_{eT}, V_{eD} and V_{eM} (the rate constants for ethanolysis of TAG, DAG and MAG, respectively). For the FFA model output the most influential parameters are V_{eEs} (the rate constant for esterification of FAAE), K_i, V_{mT}, V_{mD} *and* V_{mM} (the rate constants for hydrolysis of TAG, DAG and MAG, respectively). From the local sensitivity analysis performed by Cheirsilp and co-workers, they conclude the reaction rate increases when the parameters K_{mT}, K_{mD}, K_{mM}, K_{mF} (equilibrium constants for

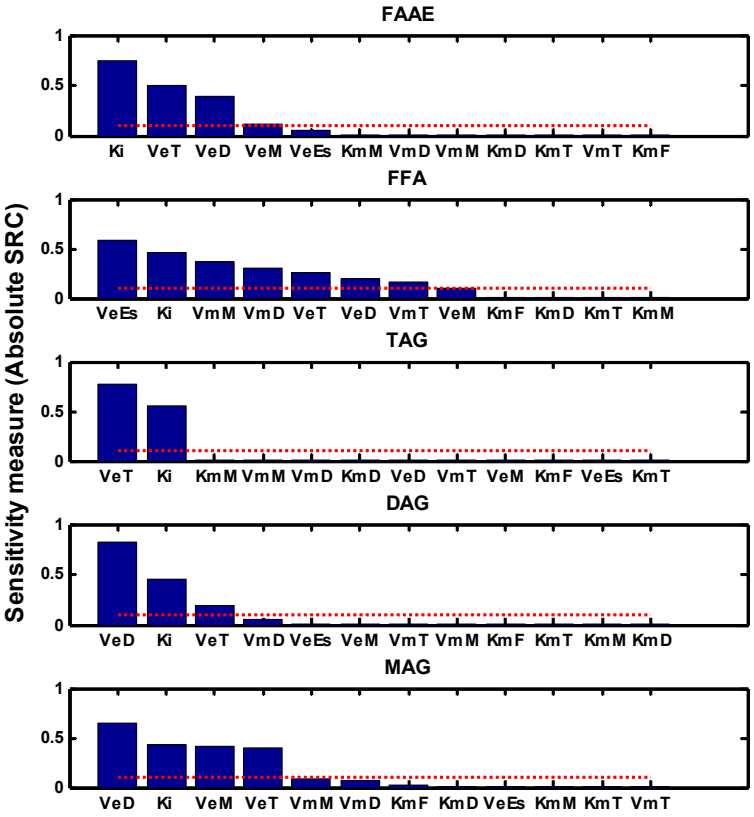


Figure 5-5 The ranked absolute values of the parameters influencing the transesterification reaction. The dashed line at 0.1 is a visual marker to show the parameters that contribute more than 1 % of the total variance in the model outputs.

TAG, DAG, MAG and FFA) decrease for the FAAE model output[6]. A different conclusion is obtained from our global sensitivity analysis. The parameters K_{mT}, K_{mD}, K_{mM}, K_{mF} are non-influential and can be fixed to any value within their ranges of uncertainty without significantly affecting the model outcomes (see section on Model simplification for the results).

A closer investigation of the 12 parameters show, K_i is the only parameter that has a significant effect on all the five model outputs. This is fully understandable given K_i is related to competitive alcohol inhibition. If the enzyme is bound to an alcohol molecule the resulting decrease in enzyme activity should affect all the model outputs.

Table 5-5 Ranked Standardized Regression Coefficients (SRC) and Estimated Means of the Distribution (μ) of Scaled Elementary Effects of the Inputs for Cheirsilp et al. Highlighted values for the SRC are those that contribute more than 1 % fraction of the total variance. The highlighted μ values are those that fall within the wedge $\mu_i = \pm 2SEM$

	FAME				FFA				TAG				DAG				MAG			
	R^2	0.95			R^2	0.94			R^2	0.83			R^2	0.96			R^2	0.91		
Rank	SRC		μ		SRC		μ		SRC		μ		SRC		μ		SRC		μ	
1	Ki	0.732	VeT	-0.616	VeEs	-0.585	VeD	-0.744	VeT	-0.760	VeD	0.562	VeD	-0.796	Ki	0.693	VeD	0.622	VeEs	-0.594
2	VeD	0.461	Ki	-0.513	Ki	0.431	Ki	-0.543	Ki	-0.513	VeM	-0.542	Ki	-0.556	VeD	0.413	VeM	-0.603	Ki	0.464
3	VeT	0.428	VeD	0.021	VmM	0.416	VmD	-0.058	VmM	-0.042	VeT	0.250	VmD	-0.069	VeT	0.334	VeT	0.314	VmM	0.384
4	VeM	0.163	VmT	-0.019	VmD	0.309	VeT	0.015	KmF	-0.020	Ki	0.222	VeT	0.023	VeM	0.139	Ki	0.221	VmD	0.295
5	VeEs	0.048	VeM	0.005	VeT	0.230	VeM	0.014	VmT	-0.018	VmM	-0.126	VmM	0.015	VeEs	0.040	VmM	-0.135	VeD	0.211
6	VmM	0.019	VeEs	0.002	VeD	0.224	VeEs	0.005	KmD	-0.005	VmD	0.049	VmT	0.004	VmD	0.010	VmD	0.068	VeT	0.166
7	VmD	0.010	KmD	0.001	VmT	0.109	VmT	0.002	KmM	-0.002	VmT	0.007	KmT	0.003	VmM	0.004	KmD	-0.014	VmT	0.109
8	KmF	0.010	KmT	0.001	VeM	-0.088	KmD	0.002	VeM	0.003	VeEs	-0.001	VeM	0.003	KmD	-0.003	KmF	0.012	VeM	-0.103
9	KmD	0.005	KmM	0.001	KmF	0.019	KmT	0.001	VmD	0.004	KmT	0.000	KmD	0.002	KmT	-0.002	VmT	-0.007	KmD	-0.002
10	KmT	-0.004	VmD	0.000	KmT	-0.019	KmM	0.001	KmT	0.005	KmM	0.000	VeEs	-0.001	KmM	-0.001	VeEs	-0.007	KmT	-0.001
11	KmM	0.003	VmM	0.000	KmD	0.005	VmM	0.001	VeEs	0.005	KmD	0.000	KmM	-0.001	VmT	0.001	KmT	-0.005	KmM	-0.001
12	VmT	0.002	KmF	0.000	KmM	0.001	KmF	0.000	VeD	0.011	KmF	0.000	KmF	-0.001	KmF	0.000	KmM	0.000	KmF	0.000
$\sum_{j=1}^{M} \beta_{j,k}$	0.961				0.920				0.843				0.948				0.921			

Morris Screening results: To easily visualize the insignificant and significant parameters on the model output, the mean, μ_i and standard deviation, σ_i of the scaled elementary effects for each model output is plotted along with the two lines formed from the standard error of the mean *Eq. (5.5)* (The standard error of the mean (SEM) gives an indication of the variability of the sample mean) .

$$\mu_i = \pm 2SEM = \pm 2\frac{\sigma_i}{\sqrt{r}} \qquad \text{Eqn. (5.5)}$$

These two lines form a wedge, if a parameter lies inside the wedge, it then indicates that the parameter has negligible effect on the model output and can be deemed non-influential. However, if the parameter lies outside the wedge, then it is said to have a significant effect on the output.

To ensure that those parameters found to be influential and those found to be non-influential maintained similar ranking, the repetition number (*r*) was incremented until the ranking of the parameters did not change. For both kinetic simulations the r value of 30 was sufficient.

From Figure 5-6, taking for example the FAAE output, V_{eT} and K_i are the most important parameters in this analysis due to their high means. V_{eT} and K_i are also the two factors with the highest standard deviation indicating the presence of non-linearity and/or interactions amongst the parameters[19](Note: A low standard deviation value, is an indication of a linear behaviour of the model for that particular parameter whereas a high standard deviation value indicates there may be nonlinear behaviour of the model for that particular parameter). The other 10 parameters are deemed non influential given they fall inside the wedge. The procedure used in analysing the FAAE output could also be applied to the other four model outputs.

The Morris Screening gives a good overview of the relative importance of uncertain factors as well as the associated non-linearity and interactions. For all of the model outputs the Morris Screening identifies a subset of factors classified as non-influential: K_{mT}, K_{mD}, K_{mM} and K_{mF} (the equilibrium constants for TAG, DAG, MAG and FFA). This agreed quite well with the Standard Regression Coefficients method. Although the ranking varies from the Standard Regression Coefficients method, the results confirm that the modeller may easily

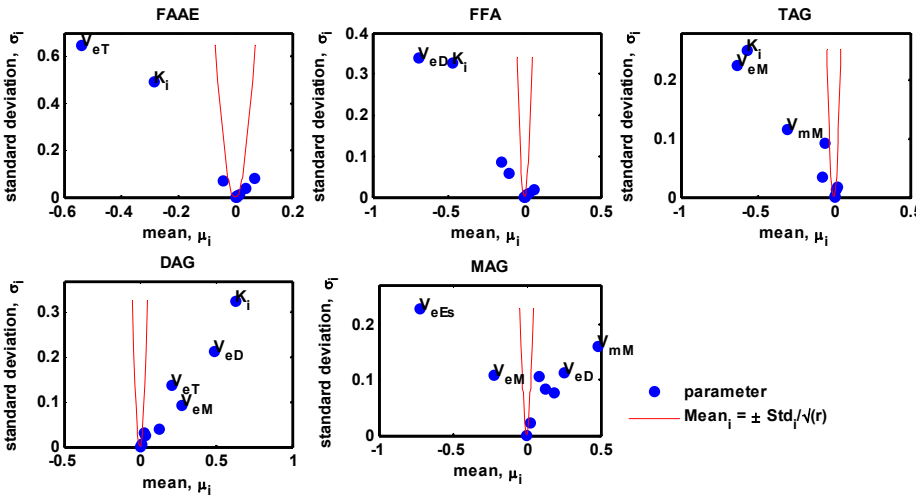


Figure 5-6 Estimated mean and standard deviation of the distribution of elementary effects of the 12 parameters for the five model outputs. For clarity only the parameters with an extreme variation in μ_i are labelled

fix the non-influential parameters. This is possible given the non-influential parameters are responsible for only a small percentage of the total output variance, thus preparing the ground for model simplification.

5.7. Comparison of Sensitivity analysis methods

Analysing the parameter ranking in Table 5-5 it is noticed that both methods give different rankings for the parameters for both models. The question then arises, which tool should be used for further work? Given the resulting coefficient of determination, R^2 is greater than 0.7 for all the model outputs; this indicates that the Standard Regression Coefficients are a valid measure of sensitivity. The bonus is that the linear regression models could also be used in place of the original model within the linear model bounds. Also the values of the Standard Regression Coefficients hold physical meaning. The sign of the coefficient indicates the effect of the parameter on the model output. Example, for the FAME model output, V_{eT} (rate constants for ethanolysis of TAG*)* has a positive Standard Regression Coefficients value of 0.428. An increase in the parameter estimate of V_{eT} will cause an increase in FAME production rate but a decrease in TAG production rate (TAG model output Standard Regression Coefficients is -0.76).

Morris Screening is found to give a good overview of the importance, interactions and non-linearity of the parameters. The method by Saltelli and co-workers, used in this paper, considers both the mean and standard deviation of the scaled elementary effects, which makes the method more resilient to identifying a factor as influential when it is not (Type I errors)[11]. However, the method can be prone to Type II errors, that is, failing to identify a factor of considerable influence on the model[11].

The two sensitivity analysis methods hence complement each other which was also found by Campolongo and co-workers[1]. It provides the modeller with more information on the parameters influence on the model outputs compared to a local sensitivity analysis. The Standard Regression Coefficients method can be used to build a linear model whose parameters represent the relative variance contribution of the parameters to the model output. The Morris Screening help to confirm the result obtained from the Standard Regression Coefficients along with highlighting non-linearity and/or interactions amongst the parameters.

5.8. Engineering Perspectives - Use of the kinetic models in enzymatic biodiesel simulation

5.8.1. Model Simplification

The non-influential parameters (K_{mT}, K_{mD}, K_{mM} *and* K_{mF}) show where research needs to be placed in devising experiments to estimate those parameters. This is quite important if recalibrating the model parameters for a different enzyme or type of substrate. For the case presented by Cheirsilp and co-workers, to aid in the model simplification, the non-influential parameters (K_{mT}, K_{mD}, K_{mM} *and* K_{mF}) were removed[6]. The parity plots Figure 5-7 show the parameters removed are essentially non-influential over the entire range of the reaction. It should be noted in the case of possible model simplification, care should be taken in removing parameters. If there is non-linearity or interactions amongst the parameters, a parameter with low importance according to ranking does not necessarily imply the factor to be non-influential. Therefore it is better to fix the value of the parameter, which is the recommendation for the equilibrium constants when recalibrating the model parameters for a different enzyme or type of substrate.

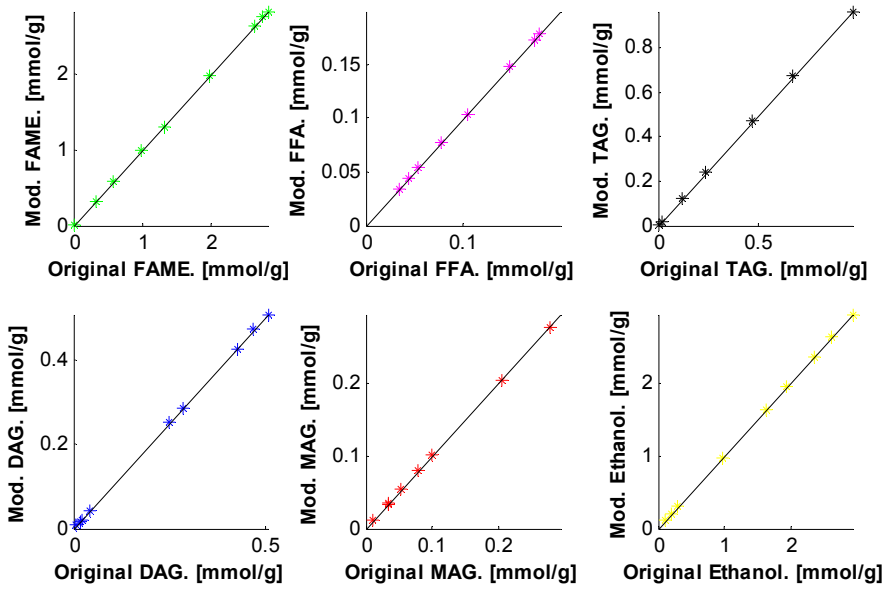


Figure 5-7 Plot showing Modified model prediction vs Original model predictions where for the modified model predictions the non-influential parameters (K_{mT}, K_{mD}, K_{mM} and K_{mF}) are removed. A R^2 value of 0.99 is obtained for all plots

5.8.2. Process Simulation

For predictive purposes such as determining when the transesterification is complete and tracking of the entire transesterification reaction the results from the uncertainty analysis showed that the parameter estimates has great potential; given the tight confidence intervals in the areas of interest. However the model in its current form is only applicable for reaction system in which the reaction goes to near complete conversion of oil to biodiesel. It should be noted that given the reaction mixture changes during the reaction (e.g. viscosity), most likely the rate constants also change during the reaction. Hence the rate constants found during the parameter estimation are just average values. This makes the uncertainty analysis a powerful tool to cater for the assumption that the parameters are fixed throughout the reaction.

For process development, an area of interest is reactor selection and configuration. A reliable kinetic model can be used to simulate and evaluate other reactor configurations, such as, fed-batch reactors and continuous stirred tank reactors (CSTR). The results from the

simulations could then be verified in the lab. One hurdle to industrial implementation of enzymatic biodiesel production is enzyme inhibition by the alcohol substrate. Substrate feeding strategies can help mitigate these inhibition effects. Simulations can be used to devise an optimal feeding policy. For fed-batch operation, during the start and end of the reaction, the model investigated should perform adequately as seen from the Monte Carlo simulations. However one should be cautious. Take for example the situation where the model is used for simulating multiple CSTR in series. It may be problematic to correlate the experimental data, with what is simulated, if operating in regions of relatively high substrate and product concentrations. Given it is in this region the model has the most uncertainty.

In this study, the focus of the sensitivity and uncertainty analysis were on the parameters of the kinetic model. The uncertainty and sensitivity analysis tools could also be used to study the effects of the process conditions. It is known the different components in the oil (TAG, DAG, FFA etc.) can vary in concentration and this could also be investigated to ascertain the effects on the model outputs. Sensitivity and uncertainty analysis on the process conditions can be an invaluable tool for the engineer in devising strategies to mitigate changes in the plant outputs due to uncertainties in the plant process conditions.

5.8.3. Process Control

One possibility is the combination of online measurements and the process model to infer the evolution of the key reaction components i.e. a soft sensor or state observer. Information from the soft sensor is used as feedback to make controlled feeding of the substrate, optimizing the process performance. In order to design the observer it is necessary to have a process model and an estimate of the noise contribution from both the model and the online measurements. The uncertainty analysis in this case provides an estimate of the noise contribution from the model due to the uncertainty in the parameter estimates. One example is in the case of using viscosity measurements. In order to monitor the progress of the transesterification reaction, Ellis and co-workers used an in-situ viscometer to correlate the viscosity measurement with the rate of biodiesel production[20]. To extend the work done by Ellis and co-workers, the rate of biodiesel production, can be coupled with the kinetic model, which can then be used to infer the concentration of the five main components measured. Use of the viscosity meter along with the soft sensor provides a relatively inexpensive way to have real time monitoring of the system; giving to

the operator quick feedback on the progress of the reaction and on the activity of the enzymes.

5.8.4. General comment on specifying input uncertainty

The confidence intervals for the parameter estimates in the work by Cheirsilp and co-workers were not presented in their work; hence a more accurate evaluation of the parameter uncertainty on the model outputs could not have been done without access to sufficient experimental data[6]. In general this highlights why when parameter estimates are being reported in literature, the confidence intervals should also be added. This will aid researchers in determining the reliability of the parameter estimates obtained along with being able to extend the uncertainty and sensitivity analysis with more rational bounds for the parameter uncertainty. The method used by Sin and co-workers in classifying the parameter input uncertainty proves useful as a starting point and engineering assumption to analyse how uncertainty in the parameters estimates influence the model outputs[2,3]. The method enabled the identification of a group of non-influential parameters which enabled simplification of the kinetic model.

Ideas brought out in this section are represented in the methodology outcomes branch in Figure 5-2.

5.9. Conclusions

A kinetic model describing the enzymatic transesterification of vegetable oil was investigated using Monte Carlo simulations of the model outputs, along with two sensitivity analysis methods based on screening and regression. The main points gleaned are:

1. The Monte Carlo simulations on the parameter estimates highlight the regions where the model is most sensitive to the uncertainty in the parameter estimates; between the 4^{th} and 25^{th} hour of the reaction. We postulated that at the start (first 4 hours) and end (last 25 hours) of the reaction the concentrations are very low, which causes the variation in the model parameters have negligible effect on the model outputs. Also the 5^{th} and 95^{th} percentile of the distribution of each model output can be used in model-based decision making such as bounds for economic process evaluation.

2. The sensitivity analysis successfully detected the influential and non-influential parameters to the model outputs. This "sets the stage" for model simplification when recalibrating the model parameters for different substrates. The non-influential parameters (K_{mT}, K_{mD}, K_{mM} and K_{mF}) can be fixed at any given value within their ranges of uncertainty without affecting significantly the model output.

The uncertainty and sensitivity analysis tools used here are used extensively throughout the rest of the thesis. During the model development, it is used to evaluate the uncertainty in the process outputs due the uncertainty in the parameter estimate (Chapter 6) as well as when an optimization is made to ascertain the uncertainty in the process outputs (Chapter 7).

5.10. List of symbols

AL	alcohol [mmol g^{-1}]
b	regression coefficient
DAG	diglycerides [mmol g^{-1}]
E*	free enzyme [g]
EE	elementary effects of the input factors
E_T	total amount of enzymes [g]
FAAE	fatty acid alkyl esters (biodiesel) [mmol g^{-1}]
FFA	free fatty acids [mmol g^{-1}]
$F_{j,k}$	distribution of the EE of the j^{th} input parameter on the k^{th} model output
G	glycerol [mmol g^{-1}]
K_i	methanol inhibition constant [mmol g^{-1}]
K_{mT}, K_{mD}, K_{mK}, K_{mF}	equilibrium constants for T, D, M, F [g $mmol^{-1}$]
M	total number of model parameters
MAG	monoglycerides [mmol g^{-1}]
N	number of measurements
N_{lh}	total number of Latin-Hypercube samples
N_t	length of the discrete time series
N_u	number of model outputs
p	number of levels used in Morris screening
R	alkyl group of the alcohol
r	number of repetitions when calculating EE
t	time [hr]
TAG	triglycerides [mmol g^{-1}]
V_{eEs}	rate constant for esterification of fatty acid ethyl ester
V_{eT}, V_{eD} and V_{eM}	rate constants for ethanolysis of TAG, DAG and MAG respectively [$mmol^{-1}$ h^{-1}]
V_{mT}, V_{mD} and V_{mM}	rate constants for hydrolysis of TAG, DAG and MAG respectively [$mmol^{-1}$ h^{-1}]
W	water [mmol g^{-1}]
y	model outputs
y_m	mean of model outputs

Greek Symbols

β	standardized regression coefficient
σ	standard deviation
Δ	perturbation factor used in the Morris screening
ε	error of the regression model
θ	model parameter
μ	mean

5.11. References

1. Campolongo F, Saltelli A. Sensitivity analysis of an environmental model: an application of different analysis methods. *Reliab Eng Syst Saf*. 1997;57:49–69.

2. Sin G, Gernaey K V, Lantz AE. Good modeling practice for PAT applications: Propagation of input uncertainty and sensitivity analysis. *Biotechnol Prog*. 2009;25:1043–1053.

3. Vangsgaard AK, Mauricio-Iglesias M, Gernaey K V, Smets BF, Sin G. Sensitivity analysis of autotrophic N removal by a granule based bioreactor: Influence of mass transfer versus microbial kinetics. *Bioresour Technol*. 2012;123:230–241.

4. Smets I, Bernaerts K, Sun J, Marchal K, Vanderleyden J, Van Impe J. Sensitivity function-based model reduction: A bacterial gene expression case study. *Biotechnol Bioeng*. 2002;80(2):195-200.

5. Kiran KL, Lakshminarayanan S. Global sensitivity analysis and model-based reactive scheduling of targeted cancer immunotherapy. *Biosystems*. 2010;101(2):117-26.

6. Cheirsilp B, H-Kittikun A, Limkatanyu S. Impact of transesterification mechanisms on the kinetic modeling of biodiesel production by immobilized lipase. *Biochem Eng J*. 2008;42:261–269.

7. Foss BA, Lohmann B, Marquardt W. A field study of the industrial modeling process. *J Process Control*. 1998;8:325–338.

8. Heitzig M, Sin G, Sales-Cruz M, Glarborg P, Gani R. Computer-Aided Modeling Framework for Efficient Model Development, Analysis, and Identification: Combustion and Reactor Modeling. *Ind Eng Chem Res*. 2011;50:5253–5265.

9. Nocedal J, Wright SJ. *Numerical Optimization*. 2nd ed. New York: Springer; 2006:1 online resource (xxii, 664.

10. Seber GAF, Wild CJ. *Nonlinear Regression*. Hoboken, NJ, USA: John Wiley & Sons, Inc; 1989.

11. Saltelli A, Ratto M, Andres T, Campolongo F, Cariboni J, Gatelli D, Saisana M, Tarantola S. *Global Sensitivity Analysis. The Primer*. Chichester, UK: John Wiley & Sons, Ltd; 2007.

12. Helton JC, Davis FJ. Latin hypercube sampling and the propagation of uncertainty in analyses of complex systems. *Reliab Eng Syst Saf*. 2003;81:23–69.

13. Al-Zuhair S. Production of Biodiesel by Lipase-Catalyzed Transesterification of Vegetable Oils: A Kinetics Study. *Biotechnol Prog*. 2005;21:1442–1448.

14. Pilarek M, Szewczyk KW. Kinetic model of 1,3-specific triacylglycerols alcoholysis catalyzed by lipases. *J Biotechnol*. 2007;127:736–744.

15. Calabrò V, Ricca E, Paola M de, Curcio S, Iorio G. Kinetics of enzymatic trans-esterification of glycerides for biodiesel production. *Bioprocess Biosyst Eng*. 2010;33:701–710.

16. Li W, Li R, Li Q, Du W, Liu D. Acyl migration and kinetics study of 1(3)-positional specific lipase of Rhizopus oryzae-catalyzed methanolysis of triglyceride for biodiesel production. *Process Biochem*. 2010;45:1888–1893.

17. Fedosov SN, Brask J, Pedersen AK, Nordblad M, Woodley JM, Xu X. Kinetic model of biodiesel production using immobilized lipase Candida antarctica lipase B. *J Mol Catal B Enzym*. 2013;85-86:156-168.

18. Fjerbaek L, Christensen K V, Norddahl B. A review of the current state of biodiesel production using enzymatic transesterification. *Biotechnol Bioeng*. 2009;102:1298–1315.

19. Campolongo F, Cariboni J, Saltelli A. An effective screening design for sensitivity analysis of large models. *Environ Model Softw*. 2007;22:1509–1518.

20. Ellis N, Guan F, Chen T, Poon C. Monitoring biodiesel production (transesterification) using in situ viscometer. *Chem Eng J*. 2008;138:200–206.

Chapter 6: Modelling of Enzymatic Biodiesel Reaction

In this chapter the development of a kinetic model for the enzymatic transesterification of rapeseed oil with methanol is presented. *A modified version of this chapter has been accepted for publication in the journal Biotechnology Progress as Price, J., Hofmann, B., Silva, V. T. L., Nordblad, M., Woodley, J. M., & Huusom, J. K. (2014). Mechanistic Modelling of Biodiesel Production using a Liquid Lipase.*

6.1. Introduction

For enzymatic biodiesel production the conventional biocatalyst is usually immobilised in order to improve enzyme recovery, which combined with increased stability, allows for re-use [1]. However, the immobilization carrier, as well as the immobilization process significantly contributes to the price of the biocatalyst, which further necessitate the re-use of the enzymes for the process to be competitive [1–3]. Within the last few years, the use of liquid lipase formulations for enzymatic biodiesel production has resulted in a significant reduction in the biocatalyst cost [4–6]. Furthermore, compared to the conventional alkali catalysts used to produce biodiesel on an industrial scale, the use of an enzymatic catalyst has the advantage that low quality feedstock's and waste oils that have a high free fatty acid (*FFA*) content can be treated. This is due to the fact that lipases are able to esterify the *FFA* contained in waste oils to esters as well as transesterify the acyl-glycerides in the oil [7]. This results in an even further reduction in the operating costs of the enzyme-catalysed biodiesel process. Nevertheless, when developing an industrial enzymatic biodiesel process a few issues related to the biocatalyst need to be addressed; methanol inhibition, deactivation at high methanol concentrations, the limited lifespan of the lipase and the relatively high cost of the enzyme per kilogram of oil treated compared to traditional chemical catalysts (e.g. sodium hydroxide) [8–10]. All of the issues outlined can be mitigated by operating the enzyme-catalysed biodiesel process in an optimal manner.

In terms of developing and optimizing the enzyme-catalysed biodiesel process efficiently, process modelling is a valuable tool to help focus the experimental work needed for process understanding and to support further process development [11]. For example, a mechanistic process model of the system can be used to simulate the process performance over a wide range of conditions and to optimize how the process is operated. Hence, a process model can further be used to develop control strategies to mitigate enzyme inactivation and improve enzyme stability. Likewise, modelling can help in the formulation of innovative process designs and configurations [12]. Integral to the development of a mechanistic model of any process, is the availability of reliable kinetic models.

Over the years, various kinetic models for the enzymatic transesterification of vegetable oils have been proposed [2,13–18]. However, only the model by Lv and co-workers address the model development for a liquid lipase formulation [2]. Likewise, in terms of engineering

design purposes such as reactor configuration, optimization and control, the gap we noticed in the literature for kinetic models for enzymatic biodiesel production is that there was a lack of emphasis placed on:

1. Use of the proposed kinetic models in the combined act of process optimization and experimental validation of the process optimization. This would enable the modeler to ascertain how well the model could be extended outside the model fitting.
2. Statistically analysing the working bounds of the model. This remains a weak point in the credibility of these kinetic models and hence their applicability for engineering design purposes.

In this chapter we address the two aforementioned points so that that we can perform model based engineering design. In using the model for engineering design, it is desired that the kinetic model describing enzymatic biodiesel production, can predict the concentration of all the major species in the reaction. It is also essential to be able to characterise how the process responds to changes in the process conditions over the entire course of the reaction for changes in:

1. **Alcohol/Oil molar ratios** - This is important given the need to balance the amount of methanol needed to shift the final equilibrium conversion of Fatty Acid Methyl Esters (FAME, biodiesel) while minimizing the effects of methanol inhibition and deactivation [1,15,19].
2. **Concentration of reactants in the reactor** - The oil composition from batch to batch can vary. The oil composition then needs to be characterised so as to ascertain when the reaction has reached within specification.
3. **Enzyme loading** - The amount of biocatalyst added will affect reaction time, enzyme efficiency and potential reuse of the enzyme.
4. **Area of the oil–water interface** - The liquid lipase Callera™ Trans L (a liquid formulation of a modified *Thermomyces lanuginosus* lipase) being used is interfacially activated and hence this phenomenon need to be taken into account when modelling the system [20,21].

None of the aforementioned models take into consideration all of the process conditions outlined for a liquid lipase formulation (As mentioned previously, the model by Lv and co-workers deal with model development for a liquid lipase formulation [2]. However, their

model formulation does not enable the evaluation of changes in enzyme concentration and the area of the oil-water interface). Hence in this work we:

- Develop and validate a dynamic mechanistic model from first principles for the transesterification of rapeseed oil with methanol using Callera™ Trans L, which takes into consideration the effects of the process conditions outlined.
- Test and experimentally validate the predictive capabilities of the model by optimizing the methanol feed profile.
- Evaluate the model structure for identifiability of the kinetic parameters (Identifiability Analysis- Correlation matrix and Collinearity index) and characterize the uncertainty in the model outputs due to the uncertainty in the parameter estimates (Uncertainty Analysis- Monte-Carlo Simulations).

This chapter is organised as follows: The proposed reaction mechanism is presented, followed by the experimental and numerical methods used. Subsequently, the results from parameter estimation and identifiability analysis are discussed. Finally, the model is used to predict an optimal methanol feeding profile and the uncertainty analysis is used to characterise the uncertainty in the model outputs.

6.2. Reaction Mechanism

6.2.1. Model formulation

The mathematical model describing the transesterification reaction in the biphasic oil–water system with a liquid lipase, Callera™ Trans L was formulated on the basis of the following assumptions:

1. The reaction proceeds via a Ping-Pong Bi-Bi mechanism [13,15,18,22]
2. Alcohol inhibition is competitive [15,18]
3. Deactivation due to the alcohol could be ignored at low methanol concentrations [15,18]
4. The interfacial and bulk concentrations of the substrate and products are the same (mass transfer from the bulk to the interface is instantaneous) [15,18]
5. Acyl migration is neglected [18]
6. All reaction steps are reversible [18,23]
7. The reactor mixture is homogenous (ideal mixing) and the density of the mixture is constant [18].

Of the various kinetic models reported [2,13–18], the mathematical formulation of the kinetic model presented by Fedosov and co-workers fulfil the first three of the process conditions outlined in the introduction section. Since their model was developed for an immobilised enzyme, it does not describe the behaviour of the liquid lipase at the oil-water interface. Hence, we extend their work by modelling the oil-water interfacial area and present a fully mechanistic formulation for the transesterification reaction using a liquid lipase. In their formulation, various pseudo-components were introduced to imitate the phase boundaries of the system, for the transesterification of rapeseed oil using Novozym 435 (immobilized Candida Antarctica lipase B) [18]. Key to how we describe the reaction mechanistically for the liquid lipase, is the interaction of the enzyme at the oil-water interface. A schematic illustration of the oil water interface along with the enzymes and its complexes is presented in Figure 6-1. The Triglycerides (*T*), Diglycerides (*D*), Monoglycerides (*M*), Biodiesel (*BD*) and Free Fatty Acid (*FA*) occupy the non-polar phase while the Bulk Enzyme (E_{bulk}), Water (*W*) Methanol (*CH*) and Glycerol (*G*) occupy the polar phase. The lipase used in this study exhibits a pronounced interfacial activation and the reaction is assumed to proceed exclusively at the interface [24]. By including the interfacial enzyme concentration (*E*), the reaction scheme proceeds as shown in Table 6-1 and Figure 6-2. The specific interfacial area per unit volume of the oil water mixture (a_T [m^2/ m^3]) can be represented as:

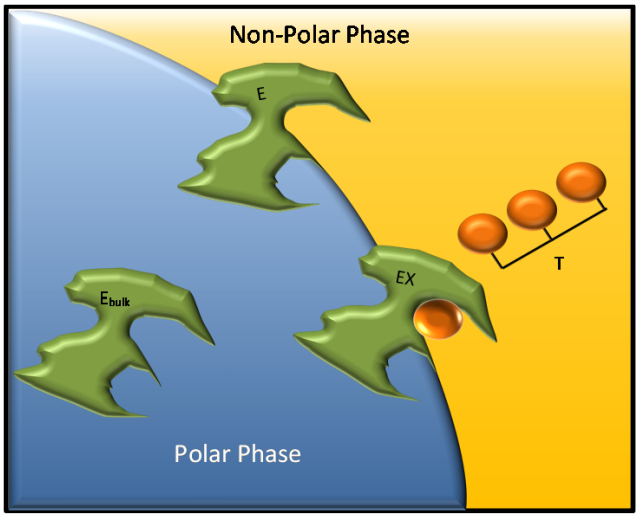


Figure 6-1 Diagrammatic representation of the enzyme at the oil water interface. The polar phase contains water, methanol, glycerol and the Free enzyme (E_{bulk}). The non-Polar phase contains the oil components along with the biodiesel formed. At the interface is the penetrated enzyme (*E*) and the Acyl Enzyme complex (*EX*)

$$a_T = \frac{6}{d_s}.\frac{V_p}{V} \quad \text{Eqn. (6.1)}$$

Where d_s is the Sauter mean diameter of the droplets in the system [m], V_p is the polar volume [m^3] and V is the bulk volume [m^3] [25,26].

An enzyme coverage A_e, of 2.98x10^7 m^2/mole, was used in this study which denotes the interfacial area that is required for the adsorption of 1 mol enzyme to the interface [20]. It is assumed that all forms of the adsorbed enzyme molecules, meaning both the free enzyme as well as all forms of the enzyme complexes and the inhibition complex, occupy the same area at the interface [20]. Given A_e, it is possible to calculate the free specific interfacial area, a_f [m^2/ m^3] as shown in equation (6.2).

$$a_f = a_T - A_e.(E + EX + ET + ED + EM + ECH) \quad \text{Eqn. (6.2)}$$

The free specific interfacial area can then be expressed as a volumetric concentration (A_f [mol/m^3]), by using the enzyme coverage to estimate a theoretical upper limit of the moles of enzyme molecules that can occupy the free interfacial area.

$$A_f = a_f / A_e \quad \text{Eqn. (6.3)}$$

In the model formulation, all concentrations used are lumped concentrations (concentration of the component in the entire reaction volume). The mass balance for the

Table 6-1 Kinetic mechanism for the enzymatic transesterification

i	Reactions		Rate of reaction (r_i)
1	$E_{bulk} + A_f \leftrightarrow E$	Enzyme in bulk absorbed at the interface	$k_1 \cdot [E_{bulk}] \cdot [A_f] - k_{-1} \cdot [E]$
2	$T + E \leftrightarrow E.T$	In reactions 2, 4 and 6 the penetrated enzyme can react with the substrate to form an enzyme substrate complex *E.T, E.D* or *E.M* (Ping)	$k_2 \cdot [T] \cdot [E] - k_{-2} \cdot [ET]$
3	$E.T \leftrightarrow EX + D$		$k_3 \cdot [ET] - k_{-3} \cdot [EX] \cdot [D]$
4	$D + E \leftrightarrow E.D$		$k_4 \cdot [D] \cdot [E] - k_{-4} \cdot [ED]$
5	$E.D \leftrightarrow EX + M$	In reactions 3, 5 and 7 the enzyme substrate complex forms the Acyl enzyme complex and releases the first product *D, M* or *G* (Pong)	$k_5 \cdot [ED] - k_{-5} \cdot [EX] \cdot [M]$
6	$M + E \leftrightarrow E.M$		$k_6 \cdot [M] \cdot [E] - k_{-6} \cdot [EM]$
7	$E.M \leftrightarrow EX + G$	Intermediate steps for the reactions were grouped together given interest is in the overall rate	$k_7 \cdot [EM] - k_{-7} \cdot [EX] \cdot [G]$
8	$EX + W \leftrightarrow FA + E$	The acyl enzyme complex can then react with water or methanol (Pong) and then release the second product *FA* or *BD* (Ping)	$k_8 \cdot [EX] \cdot [W] - k_{-8} \cdot [FA] \cdot [E]$
9	$EX + CH \leftrightarrow BD + E$		$k_9 \cdot [EX] \cdot [CH] - k_{-9} \cdot [BD] \cdot [E]$
10	$CH + E \leftrightarrow E.CH$	Reversible competitive methanol inhibition	$k_{10} \cdot [CH] \cdot [E] - k_{-10} \cdot [ECH]$

system can then be combined with the kinetics to give the system of ordinary differential equations presented below.

$$\frac{d([T].V)}{dt} = -V(r_2)$$
$$\frac{d([D].V)}{dt} = V(r_3 - r_4)$$
$$\frac{d([M].V)}{dt} = V(r_5 - r_6)$$
$$\frac{d([BD].V)}{dt} = V(r_9)$$
$$\frac{d([FA].V)}{dt} = V(r_8)$$
$$\frac{d([G].V)}{dt} = V(r_7)$$
$$\frac{d([W].V)}{dt} = -V(r_8)$$
$$\frac{d([CH].V)}{dt} = F_a CH_{feed} - V(r_9 - r_{10})$$
$$\frac{d([E].V)}{dt} = V(r_1 + r_8 + r_9 - r_2 - r_4 - r_6 - r_{10})$$
$$\frac{d([EX].V)}{dt} = V(r_3 + r_5 + r_7 - r_8 - r_9) \qquad \text{Eqn. (6.4)}$$
$$\frac{d([E.T].V)}{dt} = V(r_2 - r_3)$$
$$\frac{d([E.D].V)}{dt} = V(r_5 - r_6)$$
$$\frac{d([E.M].V)}{dt} = V(r_6 - r_7)$$
$$\frac{d([E.CH].V)}{dt} = V(r_{10})$$
$$\frac{d([E_{bulk}].V)}{dt} = -V(r_1)$$
$$\frac{d(V_p)}{dt} = R_G + R_W$$
$$\frac{d(V)}{dt} = (F_a)$$

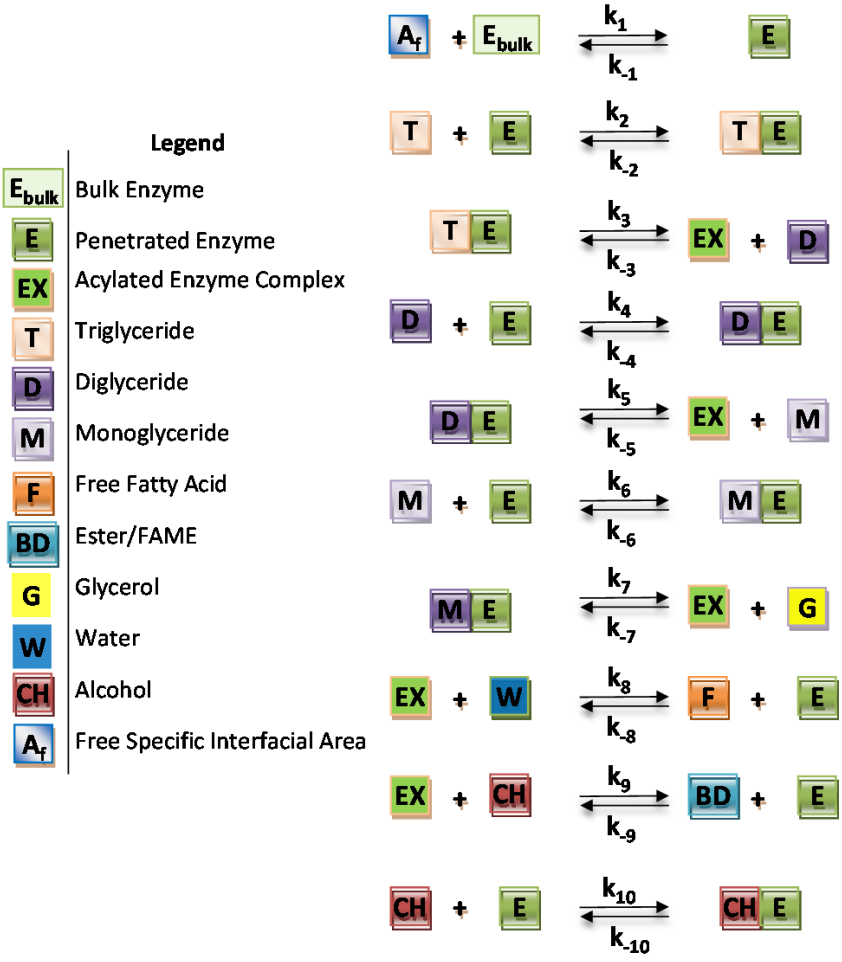


Figure 6-2 Conceptual scheme for the overall reaction mechanism describing the enzymatic transesterification

The measurement vector $y_{m,i}$ is then shown in equation (6.5) where y_m is the measurement matrix [mass %], x are the measured state variables [mol/L], V is the bulk volume and *rmm* is the relative molecular mass of component *i*.

$$y_{m,i} = \frac{x_i V\, rmm_i}{\sum_{i=1}^{5} x_i V\, rmm_i} \times 100,\ y_m = \begin{bmatrix} TAG \\ DAG \\ MAG \\ FAME \\ FFA \end{bmatrix}' \quad and\ x = \begin{bmatrix} T \\ D \\ M \\ BD \\ FA \end{bmatrix}' \qquad \text{Eqn. (6.5)}$$

6.3. Experimental Materials and Methods

6.3.1. Chemicals

Rapeseed oil was obtained from Emmelev A/S (Otterup, Denmark) and oleic acid was purchased from Sigma-Aldrich (Brøndby, Denmark). Absolute methanol (99.8%, technical grade) was purchased from VWR Bie & Berntsen A/S (Herlev, Denmark). n-Heptane (99%), acetic acid (99%), isopropanol (99%) and tert-butyl methyl ether (99.8%) for HPLC-Analysis were obtained from Sigma-Aldrich A/S (Brøndby, Denmark).

6.3.2. Biocatalyst

Callera™ Trans L with a hydrolytic activity of approximately $1x10^5$ LU/g was kindly donated by Novozymes A/S (Bagsværd, Denmark). One LU is defined as the activity required to produce 1 µmol butyric acid in the hydrolysis of tributyrin under standard conditions (pH 7.5, 0.2 M substrate) [24].

6.3.3. Fed-batch experiments

For the thirteen experiments (see Table 6-2), the water and enzyme content were varied from 3 to 7 and 0.1 to 0.5 wt. % oil respectively. In all the experiments 1.5 equivalents (Eq.) of methanol was reacted with the Rapeseed oil. One equivalent corresponds to the stoichiometric amount of alcohol needed to convert all fatty acid residues in the oil to biodiesel (i.e. 1 mol oil : 3 mol alcohol). The reaction was carried out in a 0.25 L glass reactor with a tank diameter of 55 mm (T) and 2 baffles, each 0.18×T wide. The reactor was immersed in a water bath with temperature control (Julabo Labor-technik GmbH, Seelbach, Germany) maintained at 35 °C. A rushton turbine (impeller diameter 0.44 T), spinning at 1400 rpm provided the mixing. Initially 0.2 Eq methanol was charged with the oil in the reactor. When the reaction mixture reached the reaction temperature, the amount of water and enzyme to be used in the experiment (see Table 6-2), was then added to the reactor and methanol feeding started. Methanol feeding was provided by a KNF STEPDOS .03 pump (KNF Neuberger AB, Stockholm, Sweden), calibrated prior to each experiment.

6.3.4. Sample preparation

50 µL samples were taken from the reactor and mixed with 500 µL solvent A (acetic acid and n-heptane 4:1000 v/v – mobile phase). Samples were then centrifuged at 14,500 rpm

Table 6-2. Experiments for the data fitting, validation and optimization

	Exp.	Methanol Feed Rate [Eq./h]	Initial Dose Methanol [Eq]	Water [wt.% oil]	Enzyme [wt.% oil]
Parameter Fitting	1	0.06	0.2	3	0.1
	2	0.06	0.2	3	0.2
	3	0.06	0.2	3	0.3
	4	0.06	0.2	5	0.2
	5	0.06	0.2	5	0.5
	6	0.1	0.2	5	0.3
	7	0.185 first 2hrs. 0.06 thereafter	0.2	5	0.2
	8	0.185 first 2hrs. 0.06 thereafter	0.2	5	0.5
Validation	9	0.185 first 2hrs. 0.06 thereafter	0.2	5	0.3
	10	0	0.4	5	0.5
	11	0	0.4	7	0.2
	12	0	0.4	7	0.5
Optimization	13	0.152 first 3hrs. 0.02 thereafter	0.525	5	0.5

for 5 min and 10 µL of the supernatant was mixed with 990 µL of solvent A prior to the HPLC analysis.

6.3.5. HPLC analysis

40 µl of the prepared sample was injected in the HPLC (Ultimate 3000, Dionex A/S, Hvidovre, Denmark) for analysis of triglycerides (TAGs), diglycerides (DAGs) , monoglycerides (MAGs), free fatty acids (FFAs), and fatty acid methyl esters (FAME). The separation of the different compounds was carried out with a cyanopropyl column (0.25 x 0.004 m) (Discovery ®, Cyano, Sigma Aldrich A/S, Brøndby, Denmark), U3000 auto-sampler, TCC - 3000SD column oven, U3400A quaternary pump modules and a Corona® Charged Aerosol Detector (Thermo Scientific Dionex, Chelmsford, MA, USA). A binary gradient program was employed for the separation of the different compounds using Solvent A, Solvent B (99.6% v/v tert-butyl methyl ether and 0.4% v/v acetic acid) and iso-propanol as Solvent C [27,28]. The detection of the different compounds after separation with the column was carried out by a Corona® Charged Aerosol Detector from Thermo Scientific

Dionex (Chelmsford, MA, USA) with nitrogen gas at a pressure of 241 KPa. The composition of the reaction samples was reported on a mass percentage basis, relative to the sum of quantified mass of the five analysed components (*TAG, DAG, MAG, FFA* and *FAME*). From previous experiments the standard deviation of the measurements, for *TAG, DAG, MAG, FFA* and *FAME* were found to be 0.40, 0.75, 0.18, 0.28 and 0.26 mass % respectively. Given that the oil composition can vary from different batches of oil, the initial concentration of the oil at the start of each experiment is measured. Further information on the HPLC method, HPLC accuracy and the HPLC calibration curves used in this work can be found in work previously done in our research group [29].

6.4. Numerical Methods

6.4.1. Simulation environment

The model was implemented and simulated in Matlab® (The Mathworks, Natick, MA). All the methods for performing the identifiability, sensitivity and uncertainty analysis were built on the toolbox based on the work by Sin and co-workers [30]. The following sections give further details of the methods used.

6.4.2. Parameter estimation, Confidence Intervals and Identifiability analysis

Parameter Estimation and Confidence Intervals: The 20 unknown kinetic constants (k_1-k_{10},k_{-1}-k_{-10}), were estimated by fitting the model equations to full time course data, using experiments 1 to 8 which covered the span of our operating conditions (see Table 6-2). To judge the quality of the fitting, the validation data sets were chosen so that the operating conditions fall within the parameter fitting dataset (Experiments 9) and also to evaluate how the model predicts the initial rates of the components (Experiments 10-12) . The differential equations were solved using a stiff variable order solver based on numerical differentiation formulas (*ode15s*). For the parameter fitting, the squared-sum of the relative errors between the simulated and experimental values for *TAG, DAG, MAG, FAME* and *FFA* were minimized using *fminsearch* (based on a simplex search algorithm) [31]. To quickly assess the quality of the data fitting, the histogram of residuals was used to examine the underlying statistical assumptions of the residuals having zero mean and being normally distribution. Scott's method, was used to determine the number of bins and is based upon the sample standard deviation and number of data points [32].

A robust method to calculate realistic confidence intervals for the parameter estimates $\hat{\theta}$, was made by using a bootstrap method [33,34]. To perform the bootstrap analysis, the residuals from the parameter estimates were used to generate synthetic data, which was subsequently used in a Monte-Carlo method 10,000 times to generate simulated data (Increasing the sample size from 5,000 to 10,000 did not change parameter distribution significantly). The simulated data was then used to generate a set of re-estimated parameters (matrix size [10,000 x 20]), where the mean of the distribution was used as the mean parameter estimate. The 95th and 5th percentiles of the re-estimated parameters were then used as the upper and lower bounds of the confidence intervals for the parameters estimates, respectively [35].

Identifiability analysis: The identifiability method based on the work by Brun and co-workers was used to ascertain which parameters could actually be identified from the available experimental data given the model structure [36]. The method follows three main steps. (1.) Calculation of the sensitivity matrix, (2.) scaling of the sensitivity matrix and (3.) calculation of the collinearity index for the subset of parameters.

Step 1 - Calculate sensitivity matrix S: The sensitivities of the model outputs (*T, D, M, BD* and *FA*) to the parameters were calculated by the direct differential method [37]. S has dimensions $N \times m$ (N is the number of experimental data points and m is the number of parameters). The sensitivities of the model outputs were placed in the sensitivity matrix, S:

$$S = \left.\frac{\partial x_i}{\partial \theta_j}\right|_{\hat{\theta}} \quad \text{Eqn. (6.6)}$$

Where, *i* is the total number of observations for the five model outputs for the 8 experiments, and *j*=1:*m* for each of the parameters.

Step 2 - Scaling of the sensitivity matrix, S: The non-dimensional sensitivity matrix $s_{i,j}$ and the normalised sensitivity matrix $\hat{s}_{i,j}$ were then computed:

$$s_{i,j} = S_{i,j}\frac{\Delta\theta_j}{sc_i} \text{ and } \hat{s}_{i,j} = \frac{s_{i,j}}{\|s_j\|} \quad \text{Eqn. (6.7)}$$

Where the mean estimate of θ_j is used for $\Delta\theta_j$, the mean of the experimental observations for each model output (i =1:5) is used for sc_i and $\|s_{i,j}\|$ is the Euclidean norm of the jth column of $s_{i,j}$.

Step 3 - Calculation of the Collinearity index for the different subset of parameters: The Collinearity index, γ_k for the subset of parameters *k* (*k=2:M*) is then:

$$\gamma_k = \frac{1}{\sqrt{\lambda_k}} \quad \text{Eqn. (6.8)}$$

Where λ_k is the smallest eigenvalue of $\hat{s}_k{}^T \hat{s}_k$ with $\hat{s}_k$ being a $N \times k$ sub-matrix of $\hat{s}_{i,j}$ whose columns correspond to the parameters in k. Brun and co-workers determined an empirical threshold of λ_k being below 15 which is used in this study [36].

6.4.3. Uncertainty analysis: Monte-Carlo simulations

In this work, the Monte-Carlo technique was used for the evaluation of uncertainty in the kinetic parameters, on the model outputs. This method offers global results due to the large number of model evaluations performed using the randomly sampled parameters, to obtain the distribution of the model outputs. The method samples from the input parameter space, θ and generates model outputs, *y*. The Monte-Carlo analysis of uncertainty involved three steps:

Step 1 - Specifying parameter input uncertainty: The confidence intervals of the parameters estimates, $\hat{\theta}$ were used as the upper and lower bounds of a given kinetic parameter.

Step 2 - Sampling input uncertainty: Latin hypercube sampling with correlation control was used to sample within the input parameter space given the dense stratification over the range of each sampled variable [38]. Samples were selected from the input parameter space, where each sample, θ_i contained one value for each input parameter creating a $\left[m \times N_{lh}\right]$ matrix. Where *m* stands for the total number of model parameters, and N_{lh} is the total number of Latin-Hypercube samples.

Step3 - Simulating the model using the sampling matrix: N_{lh} dynamic simulations were then performed using the $\left[m \times N_{lh}\right]$sampled input matrix. Each simulation result was stored in a $\left[N_t \times N_u \times N_{lh}\right]$ size array where, N_t is the length of the discrete time series and N_u is the number of model outputs. The complete Monte Carlo results provide a cumulative

distribution function for each output variable at each time instant. The uncertainty of the model outputs were then represented using mean and percentile calculations.

6.5. Results and Discussion

6.5.1. Parameter Estimates and Confidence intervals

The histogram of the residuals for the fitting of Exp. 1-8 were used to assess the quality of the model fitting. For the proposed model, it can be seen in Figure 6-3 that the histogram is not skewed which gives an indication that the complexity and choice of model is appropriate. Also the residuals have a mean of -0.05 mass % (approximately zero mean) and standard deviation of 2.64 mass %. This signifies that 95 % of the residuals are within -0.05±5.28 mass %. This result is reasonable, given that a mass balance on the acyl groups for the experimental data close within 3 mass %.

The parameter estimates for the proposed model are shown in Table 6-3 along with the confidence intervals and correlation matrix. Generally, the narrower the confidence interval, the higher the quality of the parameter estimate. The confidence intervals for the kinetic parameters k_{-2}, k_3 and k_{-6}, (the reverse kinetic constants for formation of the TAG enzyme substrate complex (*E.T)*, the rate of DAG production from *E.T* and the reverse kinetic constants for formation of the MAG enzyme substrate complex (*E.M)* respectively) deviate more than 40% from the mean estimates signifying a low sensitivity of the model outputs to those parameters. This may be due to the data set not having sufficient information given that the intermediate enzyme substrate complexes were not measured.

For Exp. 9-12 the histogram of the residuals (see Figure 6-4) for the validation data set is slightly skewed to the left indicating the model tends to underestimate the concentrations. The performance of the parameter estimates over the entire time course of the reaction is illustrated in Figure 6-5, using validation data set Exp.9. The proposed model captures the dynamics for the five components over the entire course of the reaction, although the prediction for *FFA* and *MAG* shows some deviation from the experimental data. We also investigated the performance of the model in the prediction of the concentrations of *TAG, DAG, MAG, FAME, FFA* and *CH* for the first 20 minutes of the reaction (see Figure 6-6), for various enzyme and water concentrations (Note: the water concentrations for Exp. 11 and 12 are outside the range used for the model fitting but are included to investigate the

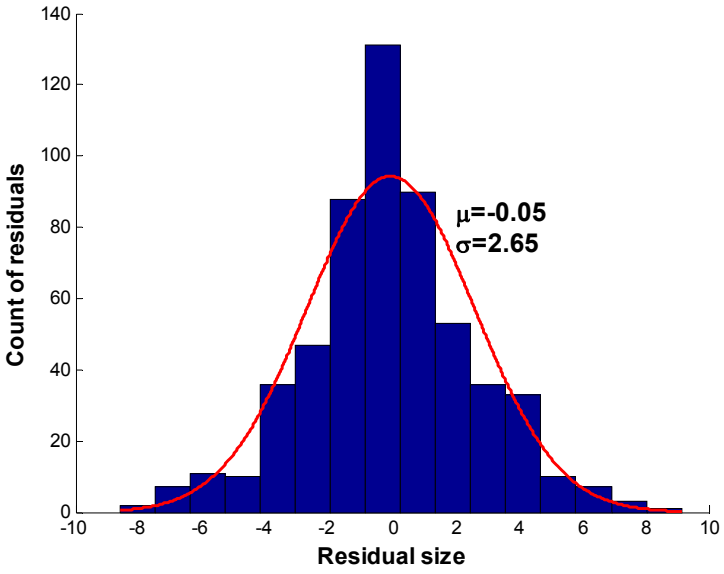


Figure 6-3 Histogram of residuals for the fitting of the proposed model for Exp. 1-8. The distribution has a mean of -0.05 mass % and a standard deviation of 2.64 mass %

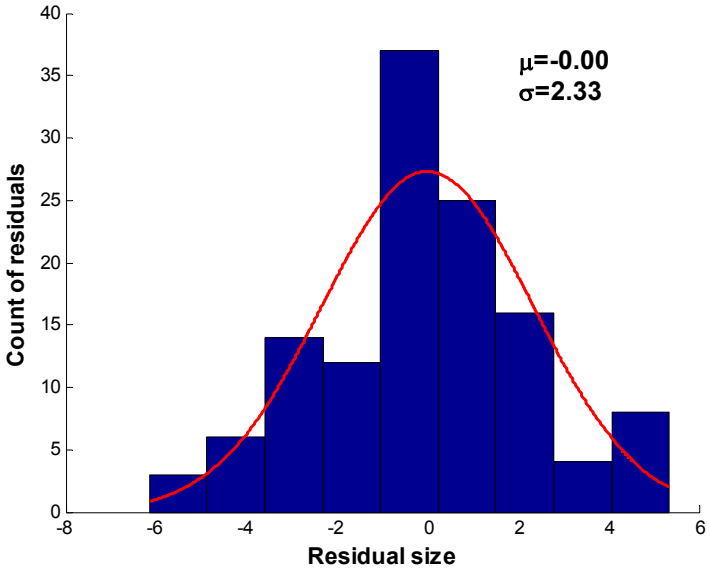


Figure 6-4 Histogram of residuals for the fitting of the proposed model for Exp. 9-12. The distribution has a mean of zero mass % and a standard deviation of 2.33 mass %

Table 6-3. Parameter estimates, related confidence intervals and correlation of the parameters for 10000 bootstrap samples. Highlighted values represent highly correlated values.

Parameter	Mean Estimate	Confidence Intervals		Identifiable parameter Subset	*Correlation Matrix*																			
		Lower	Upper		k_1	k_{-1}	k_2	k_{-2}	k_3	k_{-3}	k_4	k_{-4}	k_5	k_{-5}	k_6	k_{-6}	k_7	k_{-7}	k_8	k_{-8}	k_9	k_{-9}	k_{10}	k_{-10}
k_1 [L/mol.min]	4.95×10^{4}	3.30×10^{4}	5.88×10^{4}	-	1.00																			
k_{-1} [1/min]	6.60×10^{0}	4.26×10^{0}	7.78×10^{0}	√	0.60	1.00																		
k_2 [L/mol.min]	1.69×10^{6}	1.45×10^{6}	1.96×10^{6}	-	-0.81	-0.32	1.00																	
k_{-2} [1/min]	1.11×10^{4}	8.78×10^{3}	1.63×10^{4}	√	-0.38	-0.07	0.47	1.00																
k_3 [1/min]	2.07×10^{7}	1.53×10^{7}	3.32×10^{7}	-	-0.13	-0.15	0.16	0.51	1.00															
k_{-3} [L/mol.s]	2.20×10^{7}	1.75×10^{7}	2.51×10^{7}	-	0.21	-0.07	-0.37	-0.77	-0.20	1.00														
k_4 [L/mol.min]	3.41×10^{6}	2.68×10^{6}	4.15×10^{6}	√	-0.40	-0.32	0.46	0.18	-0.06	-0.03	1.00													
k_{-4} [1/min]	1.33×10^{7}	1.06×10^{7}	1.59×10^{7}	-	0.30	0.01	-0.52	-0.32	-0.21	0.46	-0.03	1.00												
k_5 [1/min]	1.55×10^{7}	1.36×10^{7}	1.79×10^{7}	-	-0.12	-0.01	0.24	0.40	0.25	-0.46	-0.20	-0.04	1.00											
k_{-5} [L/mol.min]	1.81×10^{5}	1.35×10^{5}	2.21×10^{5}	√	0.08	-0.02	-0.08	0.01	-0.17	0.14	0.50	-0.10	0.02	1.00										
k_6 [L/mol.min]	9.13×10^{4}	8.31×10^{4}	1.01×10^{5}	-	-0.70	-0.15	0.80	0.47	0.20	-0.52	0.26	-0.55	0.29	-0.17	1.00									
k_{-6} [1/min]	5.43×10^{5}	4.14×10^{5}	8.41×10^{5}	-	0.14	0.19	-0.04	0.20	0.12	-0.30	-0.11	-0.23	0.18	0.05	0.18	1.00								
k_7 [L/mol.min]	7.06×10^{6}	4.84×10^{6}	9.41×10^{6}	-	0.10	0.01	-0.08	0.02	0.09	0.06	-0.26	0.19	0.17	-0.16	-0.19	-0.01	1.00							
k_{-7} [L/mol.min]	4.93×10^{0}	4.45×10^{0}	5.86×10^{0}	√	0.21	-0.17	-0.40	-0.30	-0.18	0.44	-0.10	0.51	-0.27	-0.01	-0.61	-0.34	0.19	1.00						
k_8 [L/mol.min]	2.36×10^{4}	2.00×10^{4}	2.84×10^{4}	√	-0.12	-0.03	-0.09	0.00	0.07	-0.18	-0.21	-0.13	-0.02	-0.25	0.18	0.11	-0.31	-0.26	1.00					
k_{-8} [L/mol.min]	3.51×10^{6}	2.66×10^{6}	4.84×10^{6}	-	-0.54	-0.22	0.63	0.42	0.31	-0.61	0.09	-0.57	0.30	-0.35	0.72	0.12	-0.23	-0.53	0.61	1.00				
k_9 [L/mol.min]	2.54×10^{4}	2.08×10^{4}	2.85×10^{4}	√	0.04	0.02	-0.40	-0.58	-0.45	0.70	-0.06	0.49	-0.44	0.10	-0.36	-0.26	-0.06	0.41	0.01	-0.51	1.00			
k_{-9} [L/mol.min]	2.05×10^{5}	1.66×10^{5}	2.44×10^{5}	√	-0.68	-0.25	0.72	0.18	0.11	-0.29	0.18	-0.48	0.14	-0.33	0.78	-0.03	-0.16	-0.33	0.17	0.69	-0.07	1.00		
k_{10} [L/mol.min]	3.23×10^{-2}	2.69×10^{-2}	4.08×10^{-2}	√	0.03	0.18	0.15	0.40	0.27	-0.48	0.01	-0.33	0.36	0.08	0.46	0.36	0.07	-0.64	0.08	0.28	-0.45	0.04	1.00	
k_{-10} [1/ min]	4.39×10^{-4}	3.05×10^{-4}	5.08×10^{-4}	√	0.26	0.00	-0.44	-0.49	-0.26	0.54	-0.13	0.38	-0.35	0.04	-0.65	-0.24	0.00	0.69	-0.03	-0.43	0.50	-0.25	-0.81	1.00

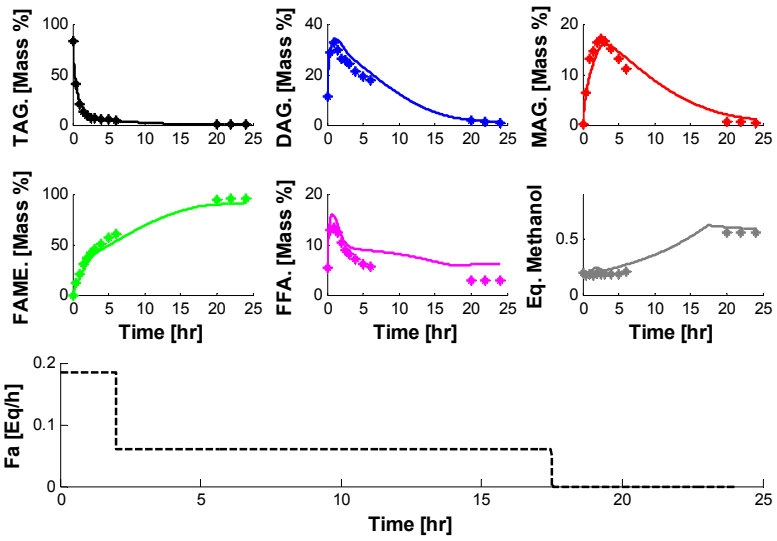


Figure 6-5 Comparison of the model fitting (-) to the Validation data set Exp 9 (*)

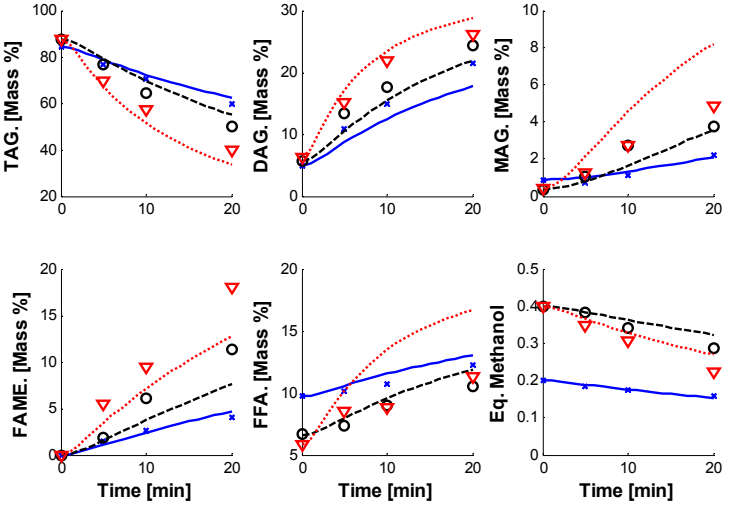


Figure 6-6. Comparison of the model fitting for the first 20 min of the reaction for experiments 10 (* experimental, - simulated), 11 (o experimental, - -simulated) and 12 (∇ experimental, … simulated).

extrapolability of the mechanistic model). The model follows the expected trends of the experimental data. For example, the *FAME* production increases, with increasing water concentration. The available interfacial area is larger which means there is an increased chance for substrates to react and hence an increase in the rate of *FAME* production [39]. Likewise, as the enzyme concentration increases from 0.2 to 0.5 wt % enzyme, the production of *FAME* increases as expected. The same reasoning extends for the other components plotted. However, at the higher water concentrations, the model tends to over predict the amount of *FFA* produced, reducing the *FAME* production.

The model mismatch observed may be due to process phenomena not taken into account. For example, the viscosity of the reaction media changes one order of magnitude over the 24 hours [40]. Hence the parameter estimates are average values of the rate constants over the entire course of the reaction. Likewise, the uncertainty in the parameter estimates plays a part given some of the parameter estimates are strongly correlated.

6.5.2. Correlated and Identifiable Parameters

Having generated parameter estimates, key to the analysis is to find out how specific the parameter estimates actually are, given the experimental data used. One simple method is to look at the correlation between the parameters (see Table 6-3). In this study ±0.75 was used to signify highly correlated parameters. This value is chosen based on the correlation value of the inhibition constant (-0.81) given the inhibition constants are usually strongly correlated [41]. For two highly correlated parameters, the change in model output due to changing one of the correlated parameters, can be compensated for by an appropriate change of the other parameter value, preventing a unique estimate of the parameter value. This can be due to the model structure or the similarity of the parameters of the underlying biological system [42]. What this may signify is that enzyme activities can be modified by changing one of the correlated parameters. For example k_{10} is highly correlated with k_{-10}. The inhibition constant has a negative coefficient for the correlation value. If one parameter value increases, the other decreases. Hence a modification of the enzyme structure that affects inhibition can potentially have a twofold effect. If the forward rate where the dead end complex *E.CH* is reduced, then the rate of disassociation of *E.CH* will increase. The

correlation seen amongst some of the parameters should be expected given the complex parallel and sequential reactions occurring simultaneously.

The plot of the Collinearity index in Figure 6-7 shows that the Collinearity index increases with the number of parameters and that only a maximum of 10 of the 20 parameters can be identified with the available experimental data. In Table 6-3 one potential subset was identified, that takes into account the parameters that are correlated. The ticked (√) parameters were the ones estimated and the others were fixed. Hence, the procedure is iterative, although in this case it only gives a reduction in the squared-sum of the relative errors between the simulated and experimental of 0.01%. It should be noted that fixing parameters, while estimating others, results in reasonable parameter values rather than " true parameter values" [36]. Given the uncertainty in the parameter estimates, we then look at how the model can be used for engineering purposes.

6.5.3. Uncertainty Analysis: Monte-Carlo Simulations

The uncertainty in the model outputs for the typically measured variables (*TAG, DAG,*

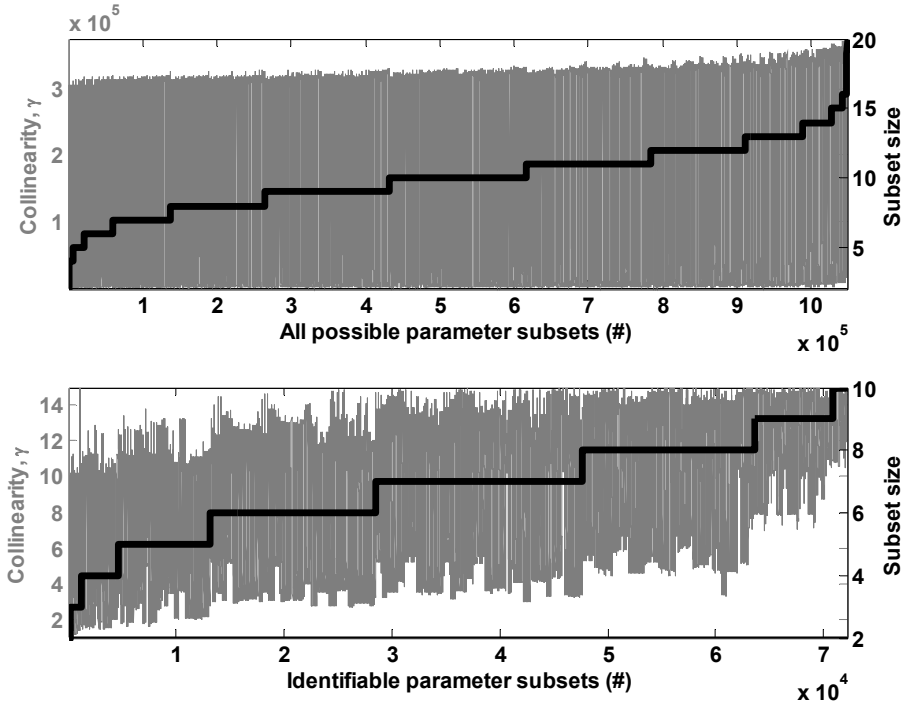


Figure 6-7. Plot of how the Collinearity index varies with all parameters (Top plot) and how the Collinearity index varies with the identifiable parameters for a threshold value of 15 (bottom plot).

MAG, FAME, FFA) during the transesterification reaction are seen in the spaghetti plots of the 500 Latin hypercube samples in Figure 6-8. When the experimental values are overlaid on the Monte Carlo simulations, we can see that over the complete course of the reaction, most of the experimental values fall within the bounds of the spaghetti plots. The narrow prediction bands for *FAME* and *TAG* reflect the robustness of the predictions for those model outputs over the entire course of the reaction, while the wide bands observed for *FFA* and *MAG* show the need for a more accurate estimate of the parameters in order to obtain more certain model predictions.

Using the cumulative frequency distribution plots (see Figure 6-9) it is possible to put bounds on the model predictions, which can give the modeller some insight into the reliability of the model to make predictions. Take for example the *FAME* predictions. At the end of the reaction the model output has a mean value of 87.9 mass % with a standard deviation of 0.64 mass %.

Given the uncertainty in the model parameters, the model gives excellent predictions of the *FAME* and *TAG* values and shows the deficiencies in the *FFA* and *DAG* predictions during the course of the reaction. We hence see this method as a valuable tool to gauge the robustness of a model to parameter uncertainty.

6.5.1. Engineering application of the model given the parameter uncertainty

For process development, a reliable kinetic model can be used, for example, to simulate and evaluate variations in feed composition, alternative reactor configurations and feeding strategies to mitigate methanol inhibition to name a few. The results from the simulations can then be tested experimentally. Below we investigate a methanol feeding strategy to mitigate inhibition and the uncertainty in the model outputs due to the parameter estimates.

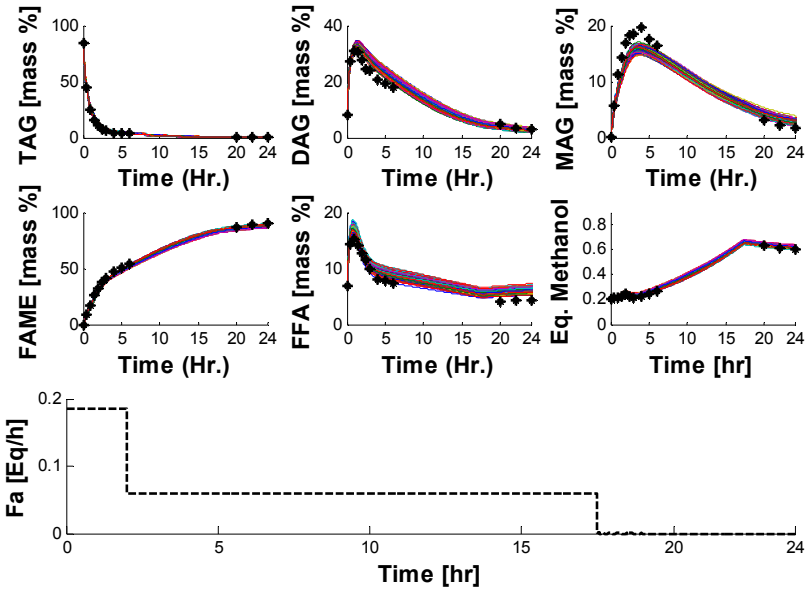


Figure 6-8. Uncertainty analysis for the validation experiment (Exp 9). The experimental values (*) are overlaid on the 500 Monte Carlo simulations (-).

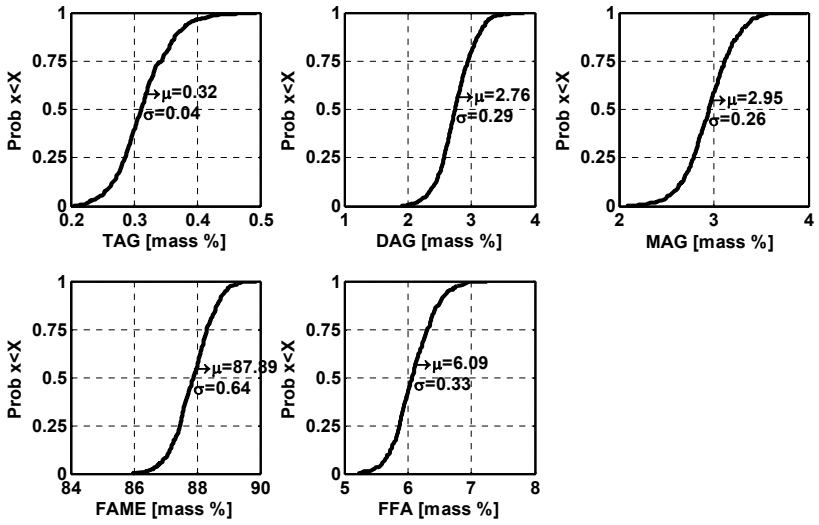


Figure 6-9 Cumulative distribution function for the 500 Monte Carlo simulations (Exp 9) at time 24 hrs

Control of methanol feeding: One hurdle to industrial implementation of enzymatic biodiesel production is inhibition and deactivation of the biocatalyst by the alcohol substrate. Simulations can be used to devise an optimal feeding policy. We followed a methanol feeding strategy similar to the one proposed by Samukawa and co-workers [43]. They found that they could increase the reuse of the immobilised enzyme (a clear indication of a reduction in enzyme deactivation), by keeping the methanol content in the reactor below the concentration that gave the highest initial rate of FAME production ($CH_{critical}$) [43]. In our previous work we found $CH_{critical}$ to be 0.525 Eq. [44]. We then combined the model of the system (equation (6.4)) with the objective function in equation (6.9). By minimizing the objective function in equation (6.9) we ensure that the methanol concentration in the reactor never goes above the critical value $CH_{critical}$ at each time step t_i by manipulating the methanol feed F_a. To simplify the experimental procedure only two step changes in the methanol feed rate were used.

$$\min_{F_a} J_{Eq} = \left(\{CH_{Eq}\}_{t_i} - \{CH_{critical}\}_{t_i} \right)^2 \quad \text{Eqn. (6.9)}$$

The optimization results are now compared to the experimental results as illustrated in Figure 6-10. The optimization objective to constrain the amount of methanol in the reactor was validated experimentally (Exp. 13) and compared to the Exp. 9 where the methanol concentration is over 0.6 Eq. for the last 7 hours of the reaction (see Figure 6-8). The optimised case respects the constraint, with the methanol concentration never going above the critical value of 0.525 Eq. for the entire reaction. However, for the optimized case, there was a 2 % reduction in the final biodiesel concentration compared to Exp. 8 which had the highest *FAME* conversion. The *FAME* equilibrium concentrations at the end of the reaction can be increased by increasing the methanol concentration at the end of the reaction. However, we are then exposing the enzyme to higher concentrations of methanol which potentially reduces the number of times the enzyme could be reused. What is interesting is the trade-off between downstream processing to bring the final biodiesel concentration within specifications and potential increase in enzyme reuse. However, this analysis is beyond the scope of this thesis.

Characterization of the model uncertainty for the optimization: The same parameter input uncertainty used for the validation data set was used for the uncertainty analysis given we are still operating within the calibrated range of the model. In Figure 6-10 the narrow prediction bands for *TAG, DAG and FAME* reflects the robustness of the predictions. When the experimental values are overlaid on the 500 Monte Carlo simulations, we can see over the course of the reaction that most of the model outputs fall within the bounds of the spaghetti plots except most notably for the *FFA* prediction. The concentrations for the *FFA* predictions are on the same order of magnitude, although the dynamics after 5 hours for the *FFA* simulation show a slight increase followed by a decrease in *FFA* concentration compared to the experimental values that show a steady decrease. As postulated earlier, given the decrease in viscosity of the reaction system as time progresses, it is believed that the rate of *FFA* consumption increases during the reaction and hence the steady decrease in the *FFA* concentration seen in the experimental data. Given the rate constants are average values in the simulation, it is not possible to capture the behaviour seen. The cumulative frequency distribution plots (see Figure 6-11) are then used to characterise the uncertainty in the model outputs. In this case, at the end of the reaction, the FAME model output has a mean value of 90.83 mass % with a standard deviation of 0.55 mass %.

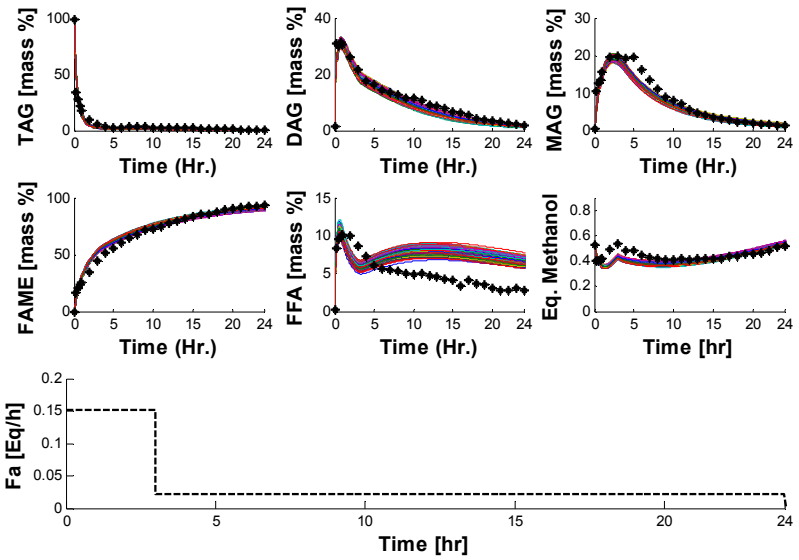


Figure 6-10 Uncertainty analysis for Exp 13. The experimental values (*) are overlaid on the 500 Monte Carlo simulations (-).

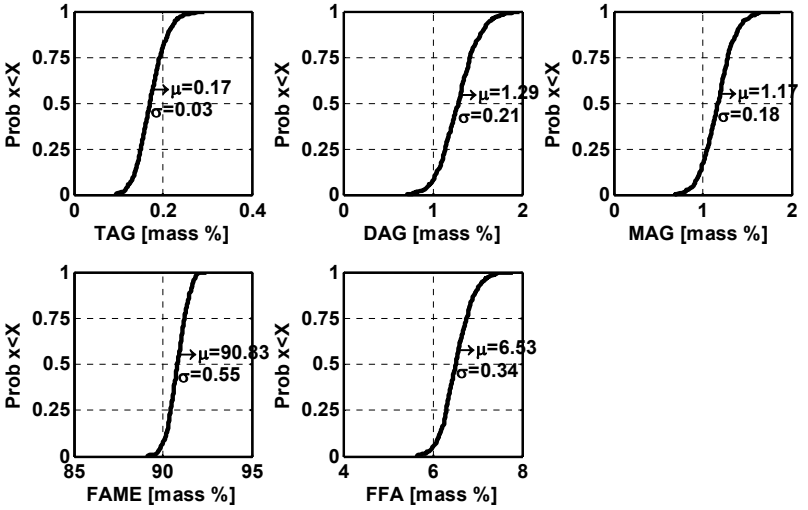


Figure 6-11 Cumulative distribution function for the 500 Monte Carlo simulations (Exp 13) at time 24 hrs.

Limitations and use of the model: The developed model is only applicable for relatively low methanol concentrations (<0.7 Eq.). To predict long term operation of the enzyme, deactivation kinetics will need to be added to the kinetic model. Also the model captures the general progress of the *FFA* produced over the course of the reaction. However, the final end point predictions are off. We mentioned earlier that the changing viscosity of the reaction media can be an issue given the reaction rates can be affected. Also water plays an important part in the hydrolysis reaction and in our model formulation bulk concentrations were used. Halling postulates that the thermodynamic water activity may be more useful for prediction of the distribution of water between the phases that is available for reaction [45]. Vantol and co-workers used thermodynamic activities instead of concentrations to model the initial rates for lipase catalysed hydrolysis [46]. The trends for the various experiments are followed quite well; however, given the increased model complexity it is still not possible to capture the experimental trends at higher substrate concentrations. From a practical point of view the model formulation based on the lumped concentrations satisfactorily captures the overall time course of the reaction for variations in the process conditions and enable us to use the model to predict and validate an optimal methanol feeding profile.

6.6. Conclusions

The present work has focused on the development of a mechanistic kinetic model, described by a system of ordinary differential equations, for the transesterification of rapeseed oil with methanol using a liquid lipase. The main purpose of the model was to capture the main effects of the process, for changes in the process conditions (alcohol/oil molar ratios, water and enzyme loadings) over the entire course of the reaction. It is clearly evident that the model formulation is "fit for purpose" when the model is used to predict, and experimentally validate an optimal methanol feeding profile so as to mitigate methanol inhibition.

The Bootstrap and the Monte-Carlo method provide realistic bounds for the parameter estimates and statistically quantify the uncertainty in the model outputs. Though these methods are computationally expensive, they provide valuable information for any model based decision making such as process design and operation. Likewise, the results give us confidence in using the developed model to evaluate and optimize enzymatic biodiesel production.

6.7. List of symbols

A_e	enzyme coverage [m^2/mole]
A_f	free specific interfacial area [mol/m^3]
a_f	free specific interfacial area [m^2/m^3]
a_T	total specific interfacial area of a droplet [m^2/m^3]
BD	biodiesel [mol/m^3]
CH, CH_{feed}	methanol in tank and in feed respectively [mol/m^3]
D	diglycerides [mol/m^3]
DAG	diglycerides [mass %]
d_s	sauter mean diameter of the droplets in the system [m]
E	interfacial enzyme concentration [mol/m^3]
E_{bulk}	total amount of enzymes in the bulk volume[mol/m^3]
Eq.	equivalents methanol
FAME	fatty acid methyl esters (biodiesel) [mass %]
FA	free fatty acids [mol/m^3]
F_a	flow rate of methanol [L/min]
FFA	free fatty acids [mass %]
G	glycerol [mol/m^3]
k_1- k_{10},k_{-1}- k_{-10}	kinetic constants for the 10 equilibrium reactions
M	monoglycerides [mol/m^3]
m	total number of model parameters
MAG	monoglycerides [mass %]

N	number of measurements
N_{lh}	total number of Latin-Hypercube samples
N_t	length of the discrete time series
N_u	number of model outputs
R_g	rate of glycerol formation [L/min]
R_w	rate of water formation [L/min]
rmm_i	relative molecular mass of component i [g/mol]
S	sensitivity matrix
$s_{i,j}$	non-dimensional sensitivity matrix
$\hat{s}_{i,j}$	normalised sensitivity matrix
t	time [min]
T	triglycerides [mol/m^3]
TAG	triglycerides [mass %]
V	bulk volume [m^3].
V_p	size of the polar volume [m^3]
W	water [mol/m^3]
y	model outputs
y_m	measurement matrix [mass %]
Greek Symbols	
γ_k	collinearity index
$\hat{\theta}$	parameter estimate

6.8. References

1. Fjerbaek L, Christensen K V, Norddahl B. A review of the current state of biodiesel production using enzymatic transesterification. *Biotechnol Bioeng*. 2009;102:1298–1315.

2. Lv D, Du W, Zhang G, Liu D. Mechanism study on NS81006-mediated methanolysis of triglyceride in oil/water biphasic system for biodiesel production. *Process Biochem*. 2010;45:446–450.

3. Tufvesson P, Lima-Ramos J, Nordblad M, Woodley JM. Guidelines and Cost Analysis for Catalyst Production in Biocatalytic Processes. *Org Process Res Dev*. 2011;15:266-274.

4. Nielsen PM, Brask J, Fjerbaek L. Enzymatic biodiesel production: Technical and economical considerations. *Eur J Lipid Sci Technol*. 2008;110:692–700.

5. Cesarini S, Diaz P, Nielsen PM. Exploring a new, soluble lipase for FAMEs production in water-containing systems using crude soybean oil as a feedstock. *Process Biochem*. 2013;48(3):484-487.

6. Pedersen AT, Nordblad M, Nielsen PM, Woodley JM. Batch production of FAEE-biodiesel using a liquid lipase formulation. *J Mol Catal B Enzym*. 2014;105:89-94.

7. Srivastava A, Prasad R. Triglycerides-based diesel fuels. *Renew Sustain Energy Rev*. 2000;4:111–133.

8. Gog A, Roman M, Toşa M, Paizs C, Irimie FD. Biodiesel production using enzymatic transesterification – Current state and perspectives. *Renew Energy*. 2012;39:10–16.

9. Al-Zuhair S. Production of biodiesel: possibilities and challenges. *Biofuels, Bioprod Biorefining*. 2007;1:57–66.

10. Adlercreutz P. Immobilisation and application of lipases in organic media. *Chem Soc Rev*. 2013;42:6406-6436.

11. Tufvesson P, Lima-Ramos J, Al Haque N, Gernaey K V, Woodley JM. Advances in the Process Development of Biocatalytic Processes. *Org Process Res Dev*. 2013;17:1233-1238.

12. Xue R, Woodley JM. Process technology for multi-enzymatic reaction systems. *Bioresour Technol*. 2012;115:183–195.

13. Al-Zuhair S. Production of Biodiesel by Lipase-Catalyzed Transesterification of Vegetable Oils: A Kinetics Study. *Biotechnol Prog*. 2005;21:1442–1448.

14. Pilarek M, Szewczyk KW. Kinetic model of 1,3-specific triacylglycerols alcoholysis catalyzed by lipases. *J Biotechnol*. 2007;127:736–744.

15. Cheirsilp B, H-Kittikun A, Limkatanyu S. Impact of transesterification mechanisms on the kinetic modeling of biodiesel production by immobilized lipase. *Biochem Eng J*. 2008;42:261–269.

16. Ricca E, Gabriela M, Stefano DP, Iorio G, Calabrò V, Paola M de, Curcio S. Kinetics of enzymatic trans-esterification of glycerides for biodiesel production. *Bioprocess Biosyst Eng*. 2010;33:701–710.

17. Li W, Li R, Li Q, Du W, Liu D. Acyl migration and kinetics study of 1(3)-positional specific lipase of Rhizopus oryzae-catalyzed methanolysis of triglyceride for biodiesel production. *Process Biochem*. 2010;45:1888–1893.

18. Fedosov SN, Brask J, Pedersen AK, Nordblad M, Woodley JM, Xu X. Kinetic model of biodiesel production using immobilized lipase Candida antarctica lipase B. *J Mol Catal B Enzym*. 2013;85-86:156-168.

19. Al-Zuhair S, Ling FW, Jun LS. Proposed kinetic mechanism of the production of biodiesel from palm oil using lipase. *Process Biochem*. 2007;42:951–960.

20. Jurado E, Camacho F, Luzón G, Fernández-Serrano M, García-Román M, Fern M, Garc M. Kinetics of the enzymatic hydrolysis of triglycerides in o/w emulsions: Study of the initial rates and the reaction time course. *Biochem Eng J*. 2008;40:473-484.

21. Hermansyah H, Wijanarko A, Kubo M, Shibasaki-Kitakawa N, Yonemoto T. Rigorous kinetic model considering positional specificity of lipase for enzymatic stepwise hydrolysis of triolein in biphasic oil-water system. *Bioprocess Biosyst Eng*. 2010;33:787-796.

22. Paiva AL, Balcão VM, Malcata FX. Kinetics and mechanisms of reactions catalyzed by immobilized lipases. *Enzyme Microb Technol*. 2000;27:187–204.

23. Fedosov SN, Xu X. Enzymatic synthesis of biodiesel from fatty acids. Kinetics of the reaction measured by fluorescent response of Nile Red. *Biochem Eng J*. 2011;56:172–183.

24. Martinelle M, Holmquist M, Hult K. On the interfacial activation of Candida-Antarctica lipase-A and lipase-B as compared with Humicola-Lanuginosa lipase. *Biochim Biophys Acta-Lipids Lipid Metab*. 1995;1258:272-276.

25. Martinez O, Wilhelm AM, Riba JP. Kinetic-study of an enzymatic liquid liquid reaction - the hydrolysis of tributyrin by Candida-Cylindracea lipase. *J Chem Technol Biotechnol*. 1992;53:373-378.

26. Al-Zuhair S, Hasan M, Ramachandran KB. Kinetics of the enzymatic hydrolysis of palm oil by lipase. *Process Biochem*. 2003;38:1155-1163.

27. Foglia TA, Jones KC. Quantitation of neutral lipid mixtures using high performance liquid chromatography with light scattering detection. *J Liq Chromatogr Relat Technol*. 1997;20:1829-1838.

28. Xu Y, Nordblad M, Woodley JM. A two-stage enzymatic ethanol-based biodiesel production in a packed bed reactor. *J Biotechnol*. 2012;162:407-414.

29. Silva VTL. Transesterification of rapeseed oil and methanol catalysed by a liquid lipase formulation in a fed-batch (Master Thesis). Instituto Superior Técnico, Lisboa, Portugal. 2013.

30. Sin G, Gernaey K V, Lantz AE. Good modeling practice for PAT applications: Propagation of input uncertainty and sensitivity analysis. *Biotechnol Prog*. 2009;25:1043–1053.

31. Lagarias JC, Reeds JA, Wright MH, Wright PE. Convergence properties of the Nelder-Mead simplex method in low dimensions. *SIAM Journal Optim*. 1998;9:112-147.

32. Scott DW. Optimal and Data-Based Histograms. *Biometrika*. 1979;66:605-610.

33. Rodriguez-Fernandez M, Egea JA, Banga JR. Novel metaheuristic for parameter estimation in nonlinear dynamic biological systems. *BMC Bioinformatics*. 2006;7:18.

34. Efron B. 1977 Rietz lecture - Bootstrap methods - Another look at the jackknife. *Ann Stat*. 1979;7:1-26.

35. Joshi M, Seidel-Morgenstern A, Kremling A. Exploiting the bootstrap method for quantifying parameter confidence intervals in dynamical systems. *Metab Eng*. 2006;8:447-455.

36. Brun R, Kuhni M, Siegrist H, Gujer W, Reichert P. Practical identifiability of ASM2d parameters - systematic selection and tuning of parameter subsets. *Water Res*. 2002;36:4113-4127.

37. Yue H, Brown M, He F, Jia JF, Kell DB. Sensitivity Analysis and Robust Experimental Design of a Signal Transduction Pathway System. *Int J Chem Kinet*. 2008;40:730-741.

38. Helton JC, Davis FJ. Latin hypercube sampling and the propagation of uncertainty in analyses of complex systems. *Reliab Eng Syst Saf*. 2003;81:23–69.

39. Hermansyah H, Kubo M, Shibasaki-kitakawa N, Yonemoto T. Mathematical model for stepwise hydrolysis of triolein using Candida rugosa lipase in biphasic oil-water system. *Biochem Eng J*. 2006;31:125-132.

40. Knothe G. Analytical methods used in the production and fuel quality assessment of biodiesel. *Trans Asae*. 2001;44:193-200.

41. Al-Haque N, Santacoloma PA, Neto W, Tufvesson P, Gani R, Woodley JM. A robust methodology for kinetic model parameter estimation for biocatalytic reactions. *Biotechnol Prog*. 2012;28:1186–1196.

42. Lovrics A, Zsely IG, Csikasz-Nagy A, Zador J, Turanyi T, Novak B. Analysis of a Budding Yeast Cell Cycle Model Using the Shapes of Local Sensitivity Functions. *Int J Chem Kinet*. 2008;40:710-720.

43. Samukawa T, Kaieda M, Matsumoto T, Ban K, Kondo A, Shimada Y, Noda H, Fukuda H. Pretreatment of immobilized Candida antarctica lipase for biodiesel fuel production from plant oil. *J Biosci Bioeng*. 2000;90:180–183.

44. Price J, Nordblad M, Woodley JM, Huusom JK. Fed-Batch Feeding Strategies for Enzymatic Biodiesel Production. In: *Proceedings of the 19th World Congress of the International Federation of Automatic Control*.; 2014:6204-6209.

45. Halling PJ. Thermodynamic predictions for biocatalysis in nonconventional media: Theory, tests, and recommendations for experimental design and analysis. *Enzyme Microb Technol*. 1994;16:178–206.

46. Vantol JBA, Jongejan JA, Duine JA, Kierkels HGT, Gelade EFT, Mosterd F, Vandertweel WJJ, Kamphuis J. Thermodynamic and kinetic-parameters of lipase-catalyzed ester hydrolysis in biphasic systems with varying organic-solvents. *Biotechnol Bioeng*. 1995;48:179-189.

PART III
Application of Modelling Tools

Chapter 7: Fed Batch Feeding Strategies

In this chapter the kinetic model developed in Chapter 6 is applied in finding an optimal methanol feeding profile where more detail is provided on the fed batch feeding strategy.

A modified version of this chapter has been accepted for publication in the Proceedings of 19th World Congress of the International Federation of Automatic Control as Price, J. A., Nordblad, M., Woodley, J., & Huusom, J. K. (2014). Fed-Batch Feeding Strategies for Enzymatic Biodiesel Production.

7.1. Introduction

Compared to the conventional alkali-catalysed biodiesel process, the enzymatic process is considered a “green reaction”. It requires less energy and is also highly selective producing a very high purity product with less downstream operations[1–3]. If the biocatalyst is to be reused, one challenge is mitigating the effects of inhibition and deactivation of the enzyme by the methanol substrate. To overcome the effects due to the methanol, researchers employ a stepwise feeding of methanol to the reactor[4–6]. However the methods that are employed are far from optimal. In order to optimize the enzymatic biodiesel process, numerous experiments are done to help characterize the system. Modelling can be a valuable tool to help focus the experimental work needed for process understanding and to support further process development. Integral to the modelling of the biodiesel process from first principles, is the availability of reliable kinetic models.

Descriptions of the various kinetic models for enzymatic transesterification of vegetable oils are quite numerous[5,7–12]. In terms of determining the optimal methanol feeding profile, the current kinetic models in literature are not able to predict the concentration of the major species over the entire course of the reaction, for changes in the process conditions such as:

- Alcohol/oil molar ratio
- Water and Free fatty acid concentrations
- Different enzyme loadings
- Interfacial area of the oil–water interface

The aim of this work is to:

- Develop a mechanistic model from first principles that takes into consideration the effects of the process conditions outlined.
- Use the proposed model to evaluate various feeding strategies to improve the biodiesel production while constraining the maximum allowable concentration of methanol in the reactor.

The chapter is organised as follows. The model formulation is presented, along with the two feeding strategies. The results of the parameter estimation are discussed along with the

results of the feeding strategies. The uncertainty in one of the feeding strategies due to the uncertainty in the parameter estimates is then investigated.

7.2. Model Formulation and Methods

7.2.1. Model formulation

The mathematical model describing the transesterification reaction in the biphasic oil–water system with a soluble lipase (Callera Trans L-*Thermomyces lanuginosus*) was formulated on the basis of the following assumptions:

1. The reaction proceeds via a Ping-Pong Bi-Bi mechanism
2. No inhibition by the substrate
3. Competitive alcohol inhibition
4. The interfacial and bulk concentrations of the substrate and products are the same (mass transfer from the bulk to the interface is instantaneous)
5. Acyl migration can be ignored
6. All reaction steps are reversible

7.2.2. Methanol Feeding Optimization

Given that the transesterification reaction is reversible, an excess of methanol is needed to push the reaction to its equilibrium conversion. For this enzyme formulation, at least 1.5 molar equivalents (Eq.) of methanol are necessary (1 Eq. of methanol corresponds to the ratio of 3 moles of methanol to 1 mole of triglyceride). However, high concentrations of methanol will cause the activity of the enzyme to decrease due to methanol inhibition and irreversibly deactivate the enzyme[13]. The mechanism for methanol inhibition is covered in the model presented, however deactivation of the enzyme is not, due to insufficient experimental data to characterise the phenomenon. Samukawa and co-workers found that they can increase the reuse of the immobilised enzyme (a clear indication of a reduction in enzyme deactivation), by using a stepwise feeding strategy. This kept the methanol content in the reactor below the concentration that gave the highest initial rate of FAME production[4]. Hence we wished to extend their work by actually being able to maintain the concentration of methanol in the reactor ($\{CH_{critical}\}$) that gave the best initial rate, at each time increment t_i, by minimizing the objective function in (1).

$$\min_{F} J_{Eq} = \left(\{CH_{Eq}\}_{t_i} - \{CH_{critical}\}_{t_i} \right)^2 \quad \text{Eqn. (7.1)}$$

The control vector for the methanol feed rate is, $F = [F_1, F_2 \ \ldots \ F_N]^T, [\mathrm{L/min}]$ and the same experimental settings in Exp. 1-7 are used along with the simulation settings in Table 7-1 to investigate the effects how the lower number of feed increments (Opt.1, N=2) and upper number of feed increments (Opt.2, N=20) affects the process.

The objective function in (2) is used to find the initial amount of methanol dosed, that achieves the highest initial rate of FAME production (IR_{FAME}). A value of 0.525 Eq. is found, and is used in the rest of the simulations.

$$\max_{CH_0} J_{IR} = IR_{FAME} \quad \text{Eqn. (7.2)}$$

7.2.3. Uncertainty analysis

The Monte Carlo method was used to propagate the uncertainty of the kinetic parameters on the output (prediction) uncertainty of the model as described in chapter 5. The confidence intervals from the parameter fitting are used to specify the input uncertainty in the parameter estimates and Latin hypercube sampling with correlation control is used for sampling of the parameters in the sample parameter space[16].

Table 7-1 Simulation settings for the Feeding strategy

Settings	**Strategy 1**	
	Opt.1	**Opt.2**
{$CH_{critical}$} [Eq.]	0.525	0.525
$CH_{@\ t=0}$ [Eq.]	0.525	0.525
Enzyme [wt.% oil]	0.5	0.5
Water [wt.% oil]	5	5
N - number of feed increments	2	20
t_{end} [min]	1500	1500

7.3. Results and Discussion

7.3.1. Feeding Strategy Simulations

The two feeding strategies simulated (Opt.1 and Opt.2), are able to satisfy the objective function in equation 7.2 at each time increment for *N=2* and *N=20*. One possible measure to ascertain which feeding strategy is better is to use the FAME yield. For the two feeding strategies simulated, it was possible to increase the FAME concentration throughout the entire course of the reaction as seen in the parity plot in Figure 7-1. Exp.7 had the highest FAME yield (703.76 g/L) of all the experiments and a reactor productivity of 28.12 g FAME L^{-1} h^{-1}. For Opt.1 and Opt.2 the increase in the FAME yield compared to Exp.7 was 4.14 % and 3.94 % respectively. What this means, from a production perspective, is that using Opt.1's feeding strategy, the reaction could be stopped 6.25 hours earlier and still have the same FAME yield as in Exp. 7. This equates to an increase in the reactor productivity of 36.9 %.

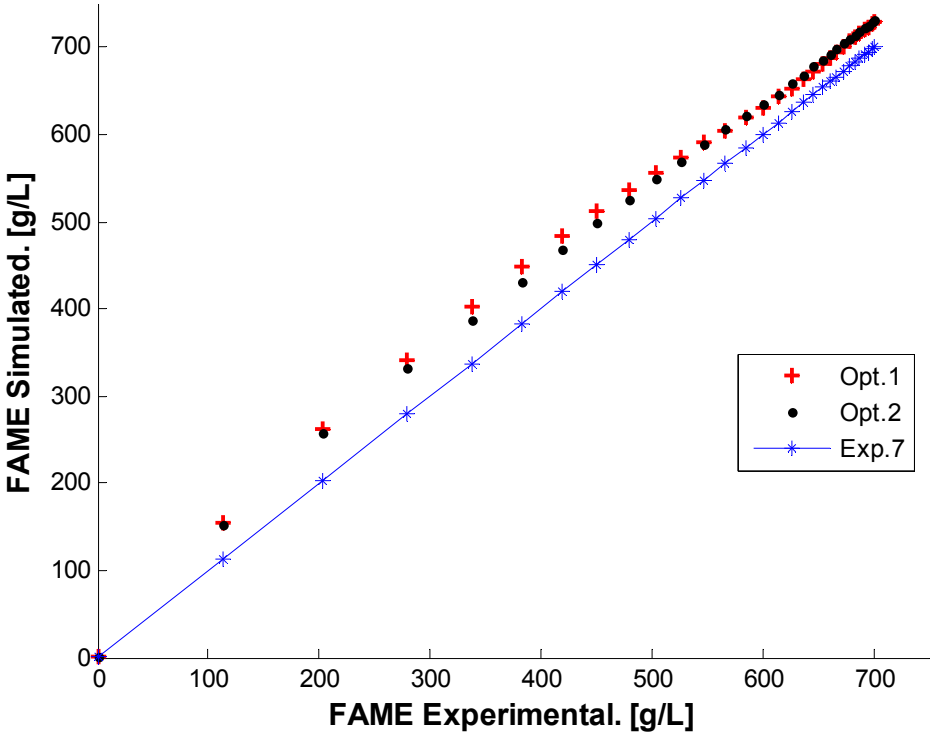


Figure 7-1 Parity plot of the Exp.7 vs. the two feeding strategies. Each point represents 50 minute increments.

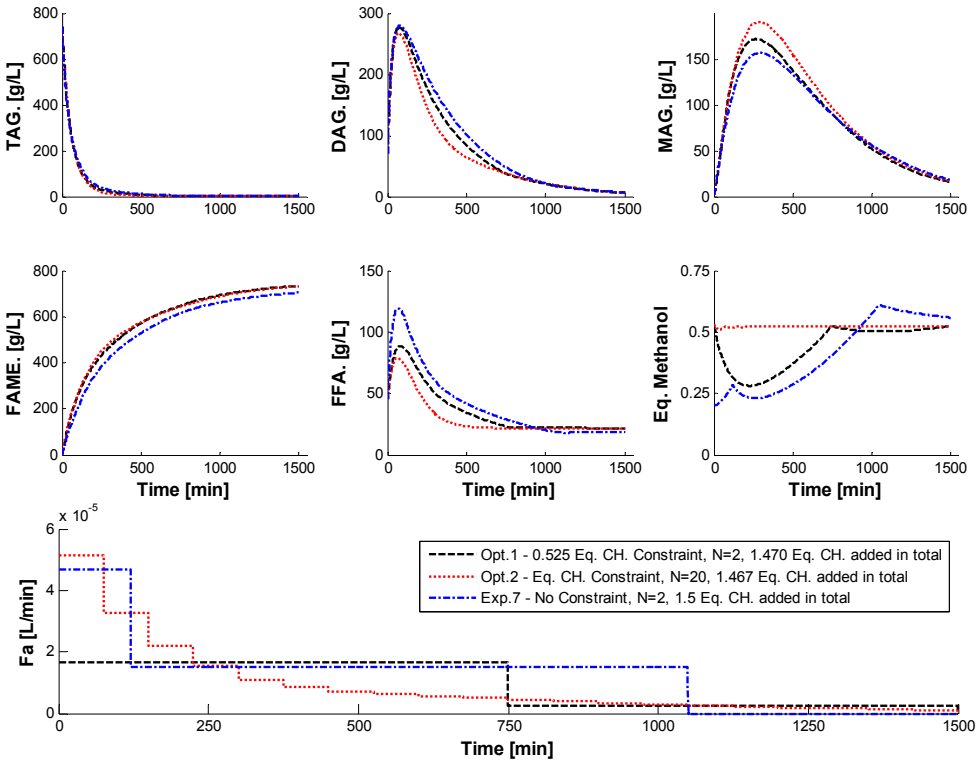


Figure 7-2 Simulation of the feeding strategies for Opt.1 and Opt. 2 along with the simulation results of Exp.7 for comparison

The increase in reactor productivity due to the optimal feeding of methanol can be explained by the plots shown in Figure 7-2. For feeding strategies Opt.1 and Opt.2 the concentration of methanol in the reactor is below or at the optimal value of 0.525 Eq. which gave the fastest initial rate. It is known that initial reaction rate increases with increasing methanol content, reaches a maximum, and thereafter decreases when the methanol content is further increased[7]. From the simulations (not shown) this behaviour also occurs during the reaction. Given the methanol concentration never crosses the critical value of 0.525 Eq. for the two feeding strategies; the inhibition is not as severe, as compared to Exp.7.

Opt.1 has the highest FAME yield in the end of the reaction compared to Opt.2 even though it does not operate at the critical FAME concentration for the entire reaction. This is due to the fact that Opt.1 is fed more methanol than Opt.2 but still less than Exp.7. This means the optimised feeding increased the biodiesel yield while decreasing the amount of

methanol that needs to be recovered in the downstream processing. The increase in FAME production for Opt.1 and Opt.2 compared to Exp.7, in the first half of the reaction is due to the increase in methanol concentration. This means there is more methanol substrate to react, giving a faster reaction before the interface is filled with other competing enzyme substrate complexes, which ultimately slows down the reaction in the later half.

Another interesting observation is that Opt.1's (also Exp.7) methanol profile for the first 700 minutes stays below 0.525 Eq. This means the enzymes in Opt.1, is not exposed to as harsh conditions as the enzymes in Opt.2 during the first half of the reaction and may provide a better environment for the enzyme, thereby decreasing the amount of enzyme that is irreversibly deactivated. However this conclusion needs to be validated in the lab by repeated reuse of the enzyme.

In Figure 7-3 we use the Monte-Carlo simulations to investigate for Opt.1, how reliable the model is given the uncertainty in the parameter estimates. The uncertainty in the model outputs is represented using the results of the Monte-Carlo simulations, obtained from the

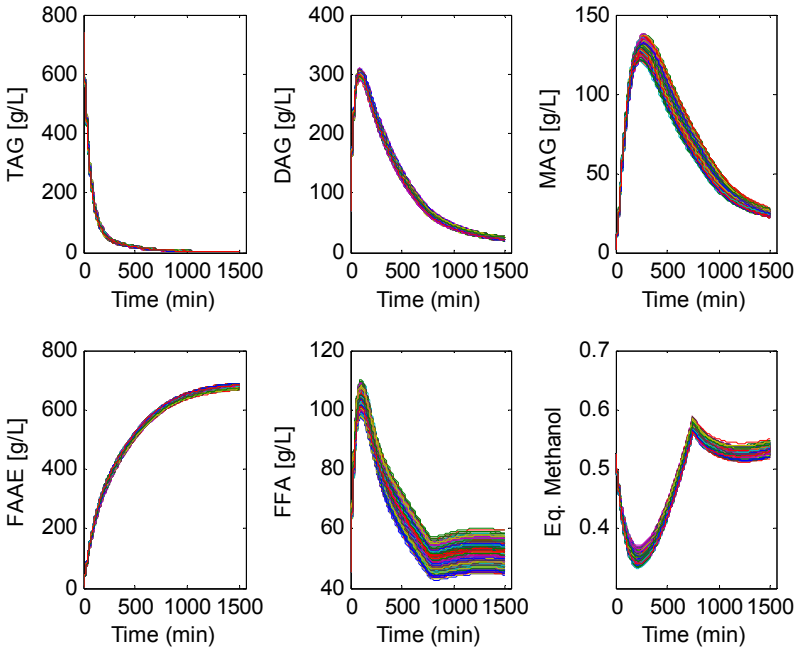


Figure 7-3 500 Monte-Carlo simulations for Opt.1 simulation depicting the uncertainty in the model predictions

dynamic simulation of the 500 Latin hypercube samples. The interpretation of the results is straightforward; the higher deviation of the 500 simulations, the worse the model prediction quality is. Overall the parameter uncertainty can be considered negligible on the model outputs even though the FFA model output shows some deviation.

7.3.2. Conclusions

The developed mechanistic kinetic model combined with the reactor mass balance enabled the evaluation of various feeding strategies to improve biodiesel production. Increasing the number of feed increments won't necessarily give a better yield but is dependent on the total amount of methanol that is feed to the reactor. It is important that the methanol concentration in the reactor is very close to the critical value to maximize the reactor productivity. In the end the two step feed feeding strategy, Opt.1 gave an increase in biodiesel yield off 4.14 %, lowered the amount of methanol that needs to be recovered and since the enzymes experiences much lower methanol concentrations this strategy may very well serve to mitigate methanol deactivation.

7.4. REFERENCES

1. Xu Y, Nordblad M, Nielsen PM, Brask J, Woodley JM. In situ visualization and effect of glycerol in lipase-catalyzed ethanolysis of rapeseed oil. *J Mol Catal B Enzym*. 2011;72:213–219.

2. Akoh CC, Chang S-W, Lee G-C, Shaw J-F. Enzymatic Approach to Biodiesel Production. *J Agric Food Chem*. 2007;55:8995–9005.

3. Nielsen PM, Brask J, Fjerbaek L. Enzymatic biodiesel production: Technical and economical considerations. *Eur J Lipid Sci Technol*. 2008;110:692–700.

4. Samukawa T, Kaieda M, Matsumoto T, Ban K, Kondo A, Shimada Y, Noda H, Fukuda H. Pretreatment of immobilized Candida antarctica lipase for biodiesel fuel production from plant oil. *J Biosci Bioeng*. 2000;90:180–183.

5. Lv D, Du W, Zhang G, Liu D. Mechanism study on NS81006-mediated methanolysis of triglyceride in oil/water biphasic system for biodiesel production. *Process Biochem*. 2010;45:446–450.

6. Du W, Xu Y-Y, Liu D-H, Li Z-B. Study on acyl migration in immobilized lipozyme TL-catalyzed transesterification of soybean oil for biodiesel production. *J Mol Catal B Enzym*. 2005;37:68–71.

7. Al-Zuhair S. Production of Biodiesel by Lipase-Catalyzed Transesterification of Vegetable Oils: A Kinetics Study. *Biotechnol Prog*. 2005;21:1442–1448.

8. Pilarek M, Szewczyk KW. Kinetic model of 1,3-specific triacylglycerols alcoholysis catalyzed by lipases. *J Biotechnol*. 2007;127:736–744.

9. Cheirsilp B, H-Kittikun A, Limkatanyu S. Impact of transesterification mechanisms on the kinetic modeling of biodiesel production by immobilized lipase. *Biochem Eng J*. 2008;42:261–269.

10. Ricca E, Gabriela M, Stefano DP, Iorio G, Calabrò V, Paola M de, Curcio S. Kinetics of enzymatic trans-esterification of glycerides for biodiesel production. *Bioprocess Biosyst Eng*. 2010;33:701–710.

11. Tan T, Lu J, Nie K, Deng L, Wang F. Biodiesel production with immobilized lipase: A review. *Biotechnol Adv*. 2010;28:628–634.

12. Fedosov SN, Brask J, Pedersen AK, Nordblad M, Woodley JM, Xu X. Kinetic model of biodiesel production using immobilized lipase Candida antarctica lipase B. *J Mol Catal B Enzym*. 2013;85-86:156-168.

13. Al-Zuhair S, Ling FW, Jun LS. Proposed kinetic mechanism of the production of biodiesel from palm oil using lipase. *Process Biochem*. 2007;42:951–960.

14. Price JA, Nordblad M, Woodley J, Huusom JK. Application of Uncertainty and Sensitivity Analysis to a Kinetic Model for Enzymatic Biodiesel Production. In: *Proceedings of the 12th IFAC Symposium on Computer Applications in Biotechnology*.; 2013:161-168.

15. Sin G, Gernaey K V, Lantz AE. Good modeling practice for PAT applications: Propagation of input uncertainty and sensitivity analysis. *Biotechnol Prog*. 2009;25:1043–1053.

16. Helton JC, Davis FJ. Latin hypercube sampling and the propagation of uncertainty in analyses of complex systems. *Reliab Eng Syst Saf*. 2003;81:23–69.

Chapter 8: From Fed Batch to CSTR Operation

Use of the developed mechanistic model to predict continuous operation in a CSTR.

8.1. Introduction

Conventionally, the industrial production of biodiesel is performed using an alkali catalyst to convert high quality vegetable oils and methanol to biodiesel in a batch reaction[1,2]. However, the main cost in the biodiesel production is the cost of the feedstock[3]. To circumvent the use of high cost feedstock's, some of industrial companies that produce biodiesel use a liquid lipase (triacylglycerol acylhydrolase, EC 3.1. 1.3) as a biocatalyst. The reason for switching to a bio-catalyst is that the use of a liquid lipase catalyst offer may advantages, such as:

- *The ability to treat a wide range of low quality/cost oil feedstocks and waste oils that have a high free fatty acid (FFA) content* - Lipases are able to esterify the *FFA* contained in waste oils to esters, as well as transesterify the acyl-glycerides in the oil, which will require additional pre-treatment steps if a conventional alkaline catalyst is used[4,5].
- *Efficient substrate utilization by the biocatalyst* - For the lipase catalysed process only 50 % excess methanol is needed to reach over 95 % biodiesel yield. An alkaline catalysed process on the other hand uses an excess of over 100% methanol which substantially increases the downstream recovery cost of the methanol.
- *Substantially lower cost of a liquid lipase compared to using an immobilised lipase*[6].
- *A higher quality glycerol by-product is produced*.[5]

When using a liquid lipase as a biocatalyst, the reaction is conventionally performed in fed-batch operation so as to minimize the inhibition and deactivation of the biocatalyst[5,7,8]. The main disadvantage of fed-batch operation is the downtime between batches. Continuous operation will afford many advantages, such as, steady state operation, smaller reactors which mean that higher mixing rates are possible and easier handling of cheaper, high melting point substrates. However, it is unclear how the continuous enzymatic biodiesel process needs to be designed and operated to ensure optimal economics. Devising a strategy for the design and operating of the process is therefore essential.

The use of conventional Levenspiel plots is an easy and effective way for sizing reactors based on batch reaction data[9]. Under certain conditions fed-batch data can also be used to guide reactor sizing for other reactor configurations, provided that the rate of change in the reactor volume is significantly smaller than the reactor volume. However, Levenspiel plots

are only valid for one reaction trajectory and are not an optimization tool. This makes the mechanistic modelling approach quite attractive. The downside to a detailed mechanistic model for a biocatalytic process is that, given the large number of parameters and the often limited experimental data points, the parameters found are not identifiable. The model is then only applicable within the operating range for which the model was validated[10]. Hence the non-identifiable parameters are fixed, while the others are estimated, resulting in reasonable parameter values rather than "true parameter values"[11].

An idea I've been contemplating on is that instead of fixing parameters what about using differences in the mass balance of the system to aid in the model fitting process. To my knowledge, I have not seen in the scientific literature (more specifically pertaining to the field of enzymatic transesterification) where the differences in the mass balance of the system are used to aid in the parameter estimation.

In this chapter, the developed mechanistic kinetic model from chapter 6 describing the transesterification reaction is used to evaluate the feasibility of a continuous process using a soluble lipase formulation. What is unique in this work compared to what is done in the scientific literature for the model calibration of kinetic models for enzymatic transesterification is the use of fed-batch and CSTR data to aid in the fitting of the model to the experimental data. Presented in this chapter is the mass balance for the system, followed by the experimental and numerical methods used for the kinetic parameter estimation. Subsequently, the results for the model validation and predictions are made along with an analysis on how the process can be operated continuously.

8.2. Process Model formulation

The same reaction mechanism is used as in Chapter 6 and is combined with the general mass balance in equation (8.1a) to give the system of ordinary differential equations which can describe the Fed-Batch and CSTR operation.

General Equation

$$\frac{dC_x}{dt} = \frac{\left(F_{o_{i-1}}C_{x_{i-1}} + F_{x_{i-1}}C_{x_{i-1}} - F_{t_i}C_{x_i} - C_x\frac{dV}{dt}\right)}{V} + r_{net_x} \quad \text{Eqn. } (8.1a)$$

With

$$\frac{dV}{dt} = F_{t_{i-i}} - F_{t_i} \quad \text{Eqn. (8.1}b\text{)}$$

Where

$$F_{t_i} = \frac{V_i}{\tau_i} \quad \text{Eqn. (8.1}c\text{)}$$

$$F_{t_i} = F_{o_{i-1}} + F_{x_{i-1}} \text{ Eqn. } (8.1d)$$

The net rates r_x can be found in Table 8-1. Where C_x is a vector of the concentration of the different components in the system, $F_{o_{i-1}}$ and $F_{x_{i-1}}$ is the volumetric flow of oil and of component x into reactor i respectively, F_{t_i} is the total volumetric flow out of reactor i, V is the working liquid volume in the reactor and τ_i is the residence time in reactor *i*.

The measurement vector $y_{m,i}$ is then shown in equation (8.2) where y_m is the measurement matrix [mass %], x_m are the corresponding measured state variables [mol/L], *V* is the bulk volume and *rmm* is the relative molecular mass of component *i*.

$$y_{m,i} = \frac{x_i V\, rmm_i}{\sum_{i=1}^{5} x_i V\, rmm_i} \times 100, y_m = \begin{bmatrix} TAG \\ DAG \\ MAG \\ FAME \\ FFA \end{bmatrix}' \; and \; x_m = \begin{bmatrix} T \\ D \\ M \\ BD \\ FA \end{bmatrix}' \quad \text{Eqn. (8.2)}$$

Table 8-1 Net rates for the various reactions

x	Net Rate of production - r_{net_x}
T	$-r_2$
D	$r_3 - r_4$
M	$r_5 - r_6$
BD	r_9
FA	r_8
G	r_7
W	$-r_8$
CH	$-(r_9 + r_{10})$
E	$r_1 + r_8 + r_9 - r_2 - r_4 - r_6 - r_{10}$
EX	$r_3 + r_5 + r_7 - r_8 - r_9$
$E.T$	$r_2 - r_3$
$E.D$	$r_5 - r_6$
$E.M$	$r_6 - r_7$
$E.CH$	r_{10}
E_{Bulk}	$-r_1$

8.3. Experimental Methods and Analysis

8.3.1. Chemicals

Rapeseed oil was obtained from a local supermarket. Absolute methanol (99.8%, technical grade) was purchased from VWR Bie & Berntsen A/S (Herlev, Denmark). n-Heptane (99%), acetic acid (99%), isopropanol (99%) and tert-butyl methyl ether (99.8%) for HPLC-Analysis were obtained from Sigma-Aldrich A/S (Brøndby, Denmark).

8.3.2. Biocatalyst

Callera™ Trans L with a hydrolytic activity of approximately $1x10^5$ LU/g was kindly donated by Novozymes A/S (Bagsværd, Denmark). One LU is defined as the activity required to produce 1 µmol butyric acid in the hydrolysis of tributyrin under standard conditions (pH 7.5, 0.2 M substrate).[12]

8.3.3. Experimental Setup

The reaction was carried out in a 2.5 L glass reactor with a tank diameter of 12 cm (T) and 5 baffles, each 0.1×T wide. Two Rushton turbines (impeller diameter 0.42 T), spinning at 515 rpm provided the mixing (power input approximately 0.6 W/L)[13]. Temperature control in the reactor was maintained at 35 °C (DT Hetotherm, Apeldoorn, Netherlands). The substrates were fed to the reactor using a KNF STEPDOS .03 pump (KNF Neuberger AB, Stockholm, Sweden), calibrated prior to each experiment. Where the oil, water-enzyme solution and methanol each had their own pump.

8.3.4. Partial fed-batch into CSTR Experiment (Fitting and model evaluation dataset)

The reactor was charged with 1980 g of oil and 0.525 equivalent (Eq.) methanol based on the oil in the reactor. One equivalent corresponds to the stoichiometric amount of alcohol needed to convert all fatty acid residues in the oil to biodiesel (i.e. 1 mol oil : 3 mol alcohol). When the reaction mixture reached the reaction temperature, the reaction was then started as a Fed-batch operation. The amount of water (5 wt%) and enzyme (0.5 wt %), was then added to the reactor and methanol feeding started (0.152 Eq./hr). After 2 hours and 20 minutes the outlet of the reactor was opened (switched to CSTR operation) and the flow rate of oil (7.52 mL/min), water-enzyme solution (0.41 mL/min) and methanol (0.47 mL/min

or 1.5 Eq./hr) to be used in the experiment, was then continuously added to the reactors with a resulting residence time of 5 hours. After steady state was reached, a step change in the methanol feed rate from 1.5 eq./hr to 3 eq./hr (based on the feed rate of oil to the reactor) was made. The reaction progress was then monitored until a new steady state was achieved.

8.3.5. Full fed-Batch Experiment (Validation data set)

The reactor was charged with 1321 g of oil and 0.2 equivalents (Eq.) methanol based on the oil in the reactor. When the reaction mixture reached the reaction temperature, the amount of water (5 wt%) and enzyme (0.5 wt %), was then added to the reactor and methanol feeding started (0.185 Eq./hr). After 2 hours the methanol feed rate was decreased to 0.06 Eq/h until 1.5 Eq of methanol was added to the reactor in total.

8.3.6. Sample preparation

Fifty-microliter samples were taken from the reactor and mixed with 500 µL Solvent A (acetic acid and n-heptane 4:1000 v/v – mobile phase). Samples were then centrifuged at 14,500 rpm for 5 min and 10 µL of the supernatant was mixed with 990 µL of solvent A prior to the HPLC analysis.

8.3.7. HPLC analysis

Forty-microliter of the prepared sample was injected in the HPLC (Ultimate 3000, Dionex A/S, Hvidovre, Denmark) for analysis of triglycerides (TAGs), diglycerides (DAGs) , monoglycerides (MAGs), free fatty acids (FFAs), and fatty acid methyl esters (FAME). The separation of the different compounds was carried out with a cyanopropyl column (0.25 x 0.004 m) (Discovery ®, Cyano, Sigma Aldrich A/S, Brøndby, Denmark), U3000 auto-sampler, TCC - 3000SD column oven and U3400A quaternary pump modules (Thermo Scientific Dionex, Chelmsford, MA, USA). A binary gradient program was employed for the separation of the different compounds using Solvent A, Solvent B (99.6% v/v tert-butyl methyl ether and 0.4% v/v acetic acid) and iso-propanol as Solvent C.[14,15] The detection of the different compounds after separation with the column was carried out by a Corona® Charged Aerosol Detector from Thermo Scientific Dionex (Chelmsford, MA, USA) with nitrogen gas at a pressure of 241 KPa. The composition of the reaction samples was reported on a mass

percentage basis, relative to the sum of quantified mass of the five analysed components (*TAG, DAG, MAG, FFA* and *FAME*).

8.4. Numerical Methods

8.4.1. Model calibration

The 20 unknown kinetic constants (k_1-k_{10},k_{-1}-k_{-10}), were estimated by fitting the model equations with the partial fed-batch into CSTR experimental data. Which was comprised of the partial fed-batch portion of the reaction (first 2 hrs and 20 min of the reaction) and the initial CSTR portion of the reaction where the methanol feed rate was 1.5 Eq/hr. This is illustrated as the grey and red sections in Figure 8-1 using the FAME concentration profile.

8.4.2. Model evaluation and validation

To judge the quality of the fitting, the step change portion of the CSTR reaction where the methanol feed rate was changed to 3 Eq/hr was used (green section in Figure 8-1). The full fed-batch data was used as the validation data set to evaluate how well the model was able to fit fed-batch where the reaction goes to completion.

8.4.3. Parameter estimation and Confidence Intervals

The differential equations were solved using a stiff variable order solver based on numerical differentiation formulas (*ode15s*). For the parameter fitting, the squared-sum of

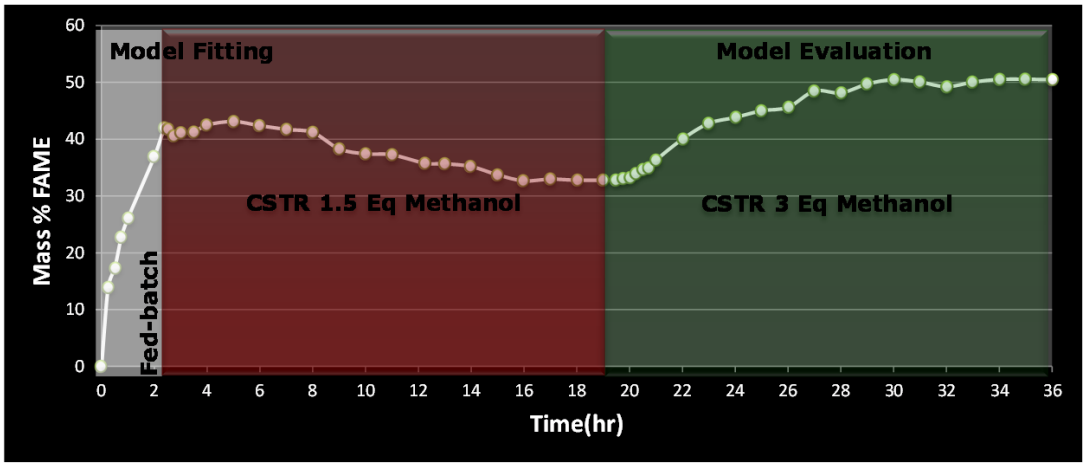


Figure 8-1 Illustration using the FAME concentration profile of the portions of the single experiment that are used for model fitting (grey and red) and the portions of the experiment used for the evaluation of the models predictive properties (green).

the relative errors between the simulated and experimental values for *TAG, DAG, MAG, FAME* and *FFA* were minimized using *fminsearch* (based on a simplex search algorithm)[28]. To quickly assess the quality of the data fitting, the histogram of residuals are used to examine the underlying statistical assumptions of the residuals having zero mean and being normally distribution. Scott's method, is used to determine the number of bins and is based upon the sample standard deviation and number of data points [29].

Using the bootstrap method, 5,000 bootstrap samples were used to estimate the confidence interval of the parameter, where the mean of the distribution is used as the mean parameter estimate. The 95^{th} and 5^{th} percentiles of the re-estimated parameters were then used as the upper and lower bounds of the parameter estimates, respectively [32].

8.5. Results and Discussion

8.5.1. Parameter Estimates and Confidence intervals of the parameters

The model captures the dynamics for the five components over the entire course of the reaction for the three different stages of the reaction as seen in Figure 8-2. The previous fitting of the model which was done on only Fed-Batch data in chapter 6 is also shown. The combined partial fed-Batch and CSTR experiment fitting has much smaller residuals compared to the previous fitting for all the measured components. The model qualitatively follows the model evaluation part of the dataset (after 19.5 hours) and gives good predictions for the endpoint value compared to using the previous kinetic constants determined in chapter 6. This is important to note given that being able to predict the concentration of the acylglycerides and FFA at the end of the reaction is just as crucial as the FAME concentration given that a product specification has to be met. Also shown is the fitting of the validation dataset (Figure 8-3) which is based on the full fed-batch data only. The newly determined kinetic constants also fit this dataset quite well and give comparable performance to the previously determined kinetic constants (see chapter 6).

To evaluate the quality of the model fitting, the histogram of residuals is shown in Figure 8-4 for the model fitting and model evaluation data given the small dataset. The histogram is slightly skewed to the left indicating the model under-predicts the values of some of the components. However, given the small number of data points (245) the skewedness is

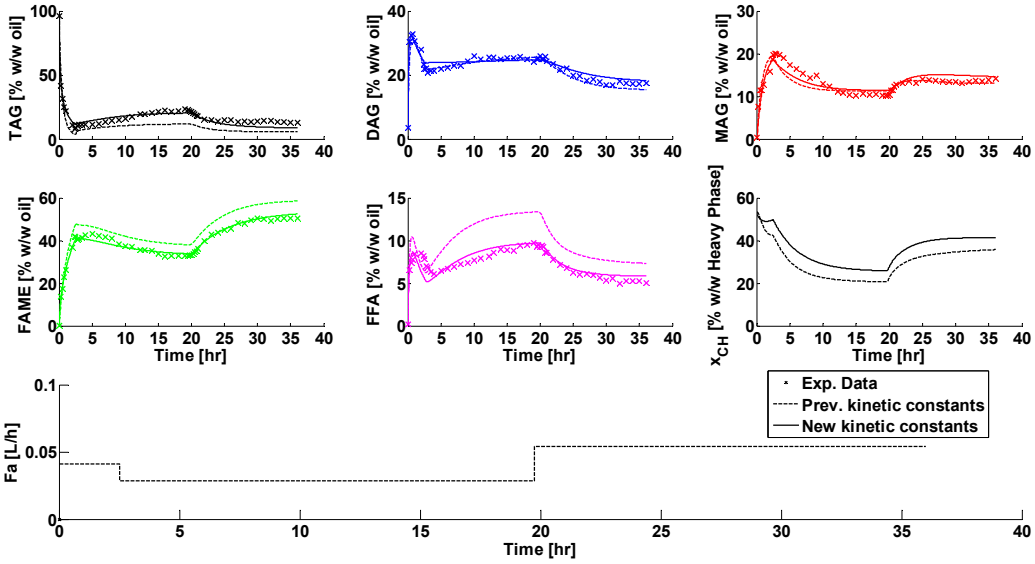


Figure 8-2 Comparison of the previously reported kinetic constants to the new kinetic constants for the fitting of the combined Fed-batch and CSTR experimental data. Note the first 19.5 hours are used for fitting and the rest of the data is used to evaluate the predictive qualities of the model.

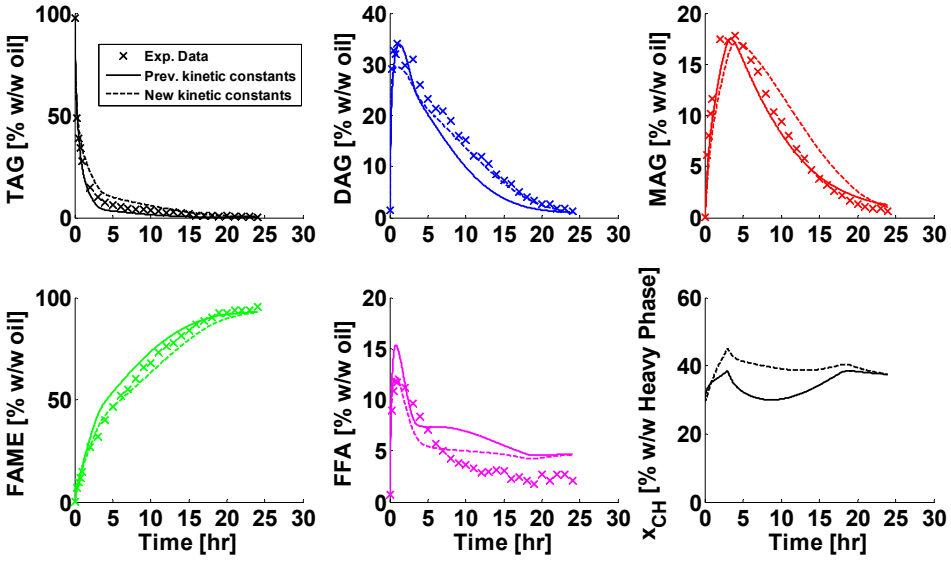


Figure 8-3 Comparison of the previously reported kinetic constants to the new kinetic constants on the Fed-Batch validation dataset(0.5 % (wt. Enzyme /wt. Oil), 0.5 % (wt. Water /wt. Oil) and feeding 1.5 times the stoichiometric amount of methanol in total over 24 hrs.

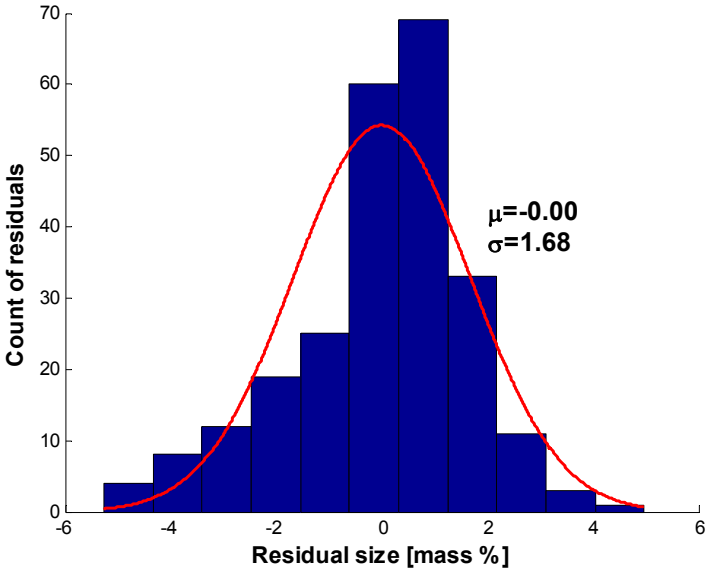


Figure 8-4 Histogram of residuals (ysim-y_m) for the partial Fed-batch into CSTR dataset using the new parameter estimates. The distribution has a mean zero mean and a standard deviation of 1.68 mass %.

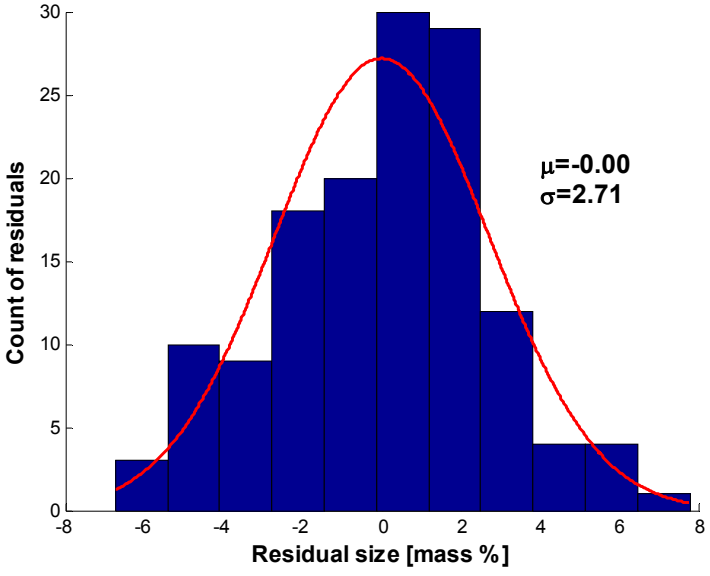


Figure 8-5 Histogram of residuals (ysim-y_m) for the validation dataset using the new parameter estimates. The distribution has a mean zero mean and a standard deviation of 2.71 mass %.

reasonable and the histogram gives an indication that the complexity and choice of model is appropriate. Given a mass balance on the acyl groups for the experimental data close within 3 mass % the standard deviation of the residuals are reasonable. The histogram for the model validation data set is shown in Figure 8-5 and given the smaller number of data point (140), the skewedness is deemed reasonable.

The parameter estimates along with the confidence intervals and correlation matrix are shown in Table 8-2. The confidence intervals along with the correlation matrix also give an indication of the quality of the parameter estimates. Generally, the narrower the confidence interval, the higher the quality of the parameter estimate. Most of the parameters have quite reasonable confidence intervals except for the inhibition constants (k_{10} and k_{-10}) whose range compared to the mean parameter estimate is over 100%. However, what is excellent is that the strong correlation between most of the parameters have been reduced compared to the correlation for the parameter estimates in chapter 6 which showed 5 pairs of highly correlated parameters (a correlation coefficient of ±0.75 was used to signify highly correlated parameters)[16]. Usually to arrive at better parameter estimates and reduce correlation between parameters various experiments are performed at different experimental conditions (e.g. variance in enzyme and methanol concentrations). To my knowledge this is the first time that it is shown that by using differences in the reactor mass balance (Fed-batch into CSTR operation) that the correlation between the parameters are reduced. If this is a coincidence or not still needs to be evaluated.

8.5.2. Reactor Simulations

Number of CSTR's and conditions to achieve comparable Fed-Batch performance: The model is now used to investigate the number of tanks and operating conditions to achieve the same biodiesel conversion as the best fed-Batch experiment (see chapter 6). In that experiment the final biodiesel value after 24 hours (t_{batch}) was 95.6 mass % using 0.5 wt% enzyme 5 wt% water and 1.5 Eq. of methanol fed in total. Assuming a 6 hour emptying and filling time (t_{ef}) for a 300 m^3 reactor (liquid volume) to achieve a similar productivity (g Fame.L.h^{-1}) in a CSTR as the Fed-Batch operation will require the total residence time in the CSTR to be less than 30 hours (t_{batch}+t_{ef}).

Table 8-2 Mean Parameter estimates, related confidence intervals and correlation of the parameters from the 5000 bootstrap samples.

Parameter	Mean Estimate	Confidence Intervals		Correlation Matrix																			
		Lower	Upper	k_1	k_{-1}	k_2	k_{-2}	k_3	k_{-3}	k_4	k_{-4}	k_5	k_{-5}	k_6	k_{-6}	k_7	k_{-7}	k_8	k_{-8}	k_9	k_{-9}	k_{10}	k_{-10}
k_1 [L/mol.min]	1.32×10^5	1.14×10^5	1.43×10^5	1.00																			
k_{-1} [1/min]	4.35×10^0	4.17×10^0	4.49×10^0	0.06	1.00																		
k_2 [L/mol.min]	1.66×10^6	1.23×10^6	1.85×10^6	0.33	-0.11	1.00																	
k_{-2} [1/min]	7.15×10^4	6.21×10^4	7.54×10^4	0.17	-0.09	0.64	1.00																
k_3 [1/min]	1.46×10^6	1.36×10^6	1.66×10^6	-0.17	-0.02	-0.64	-0.25	1.00															
k_{-3} [L/mol.min]	5.58×10^7	5.06×10^7	6.36×10^7	-0.02	-0.35	0.25	-0.02	0.10	1.00														
k_4 [L/mol.min]	6.84×10^6	5.71×10^6	7.32×10^6	0.32	0.06	0.54	0.29	-0.40	-0.30	1.00													
k_{-4} [1/min]	7.10×10^5	6.74×10^5	7.86×10^5	-0.32	-0.15	-0.39	-0.33	0.47	0.28	-0.37	1.00												
k_5 [1/min]	4.21×10^4	3.95×10^4	4.75×10^4	-0.25	-0.17	-0.25	-0.22	0.40	0.39	-0.66	0.50	1.00											
k_{-5} [L/mol.min]	1.42×10^6	1.24×10^6	1.59×10^6	0.09	-0.17	0.31	0.46	-0.25	0.14	0.23	-0.52	0.02	1.00										
k_6 [L/mol.min]	4.39×10^4	3.93×10^4	4.64×10^4	0.04	0.13	0.36	0.19	-0.38	-0.15	0.43	-0.34	-0.47	0.01	1.00									
k_{-6} [1/min]	1.02×10^7	9.09×10^6	1.11×10^7	0.05	-0.05	0.10	0.06	-0.05	-0.02	0.10	0.00	-0.21	-0.22	0.60	1.00								
k_7 [L/mol.min]	4.14×10^6	3.95×10^6	4.37×10^6	-0.23	0.08	-0.04	-0.14	0.03	0.06	-0.06	0.13	0.09	-0.15	-0.27	0.26	1.00							
k_{-7} [L/mol.min]	2.91×10^0	2.74×10^0	2.96×10^0	0.10	0.21	0.03	0.11	0.02	-0.06	0.01	-0.02	-0.03	-0.02	0.07	0.06	0.00	1.00						
k_8 [L/mol.min]	1.39×10^5	1.26×10^5	1.58×10^5	-0.32	0.12	-0.63	-0.38	0.36	0.07	-0.57	0.25	0.23	-0.16	-0.24	-0.04	0.00	-0.02	1.00					
k_{-8} [L/mol.min]	3.62×10^5	3.28×10^5	3.83×10^5	0.11	0.29	-0.16	-0.33	0.08	-0.31	0.07	-0.09	-0.09	-0.42	0.14	0.10	0.05	-0.01	0.44	1.00				
k_9 [L/mol.min]	3.29×10^5	3.08×10^5	3.75×10^5	-0.44	0.06	-0.58	-0.49	0.23	0.23	-0.63	0.32	0.33	-0.10	-0.34	-0.14	0.07	-0.01	0.64	-0.09	1.00			
k_{-9} [L/mol.min]	4.61×10^4	4.33×10^4	5.14×10^4	-0.05	0.19	-0.47	-0.68	0.29	-0.16	-0.30	0.24	0.29	-0.53	-0.25	-0.13	0.18	0.01	0.21	0.41	0.44	1.00		
k_{10} [L/mol.min]	2.73×10^3	1.51×10^3	7.94×10^3	0.04	0.08	0.06	0.13	-0.04	-0.04	0.03	-0.15	-0.07	0.10	0.13	0.04	-0.10	0.20	-0.05	-0.04	-0.05	-0.11	1.00	
k_{-10} [1/min]	3.07×10^5	2.47×10^4	4.06×10^5	0.07	0.20	-0.05	-0.13	0.07	0.00	0.01	0.03	0.09	-0.03	-0.05	-0.03	0.09	-0.01	-0.01	0.09	-0.01	0.12	0.34	1.00

Given the kinetic model does not have enzyme deactivation kinetics; the concentration of methanol in the heavy phase (heavy phase is composed of water, methanol and glycerol) is kept below 42 mass % methanol in the CSTR simulations. This was the amount of methanol in the heavy phase in the combined partial fed-Batch/CSTR experiment during the latter half of the reaction, where the methanol feed rate was increased to 3 Eq. methanol based on the oil flow rate (see Figure 8-2). During this period the FAME production did not decrease over the 16.5 hours and it was assumed that the enzyme kept most of its activity.

The simulations show that with 5 tanks and a residence time of 6 hours in each reactor a FAME conversion of 93.9 mass % is achieved as seen in Figure 8-6. To achieve 95.6 mass % FAME, the overall residence time will need to be increased to 40 hours. Given the reaction slows down in the latter half of the reaction this indicates that the latter half of the reaction should be performed in a plug flow or Fed-Batch reactor. To be able to use off the shelf equipment and still have the flexibility of a continuous process, the use of a combination of CSTR and Fed-Batch reactors to perform the transesterification reaction is investigated.

Alternative configuration: CSTR into Fed-Batch: For the alterative configuration, having the CSTR up front means that the converted FAME in the first CSTR can solubilise higher melting point substrates that are cheaper than virgin oils e.g animal fats. The CSTR outlet is fed into the fed-Batch reactor to finish the reaction. By having multiple Fed-Batch reactors they can be switched to keep continuous operation.

For the simulation it is assumed that the first reactor is operated as a CSTR with the same conditions as the third part of the CSTR experiment where the methanol feed rate is stepped up to 3 Eq/hr. Two scenarios are investigated. The last reactor is operated as a Fed-batch where additional methanol is continually feed to the end of the reaction or as a batch where no methanol is added once the reactor is full. As seen in Figure 8-7 methanol needs to be added to ensure a high biodiesel yield and reduce the FFA concentration. However, given the high methanol content in the heavy phase compared to the batch operation the methanol recovery is more process intensive. At 25 hours the FAME yield is 94.5 mass % which then gives an overall residence time of 30 hours.

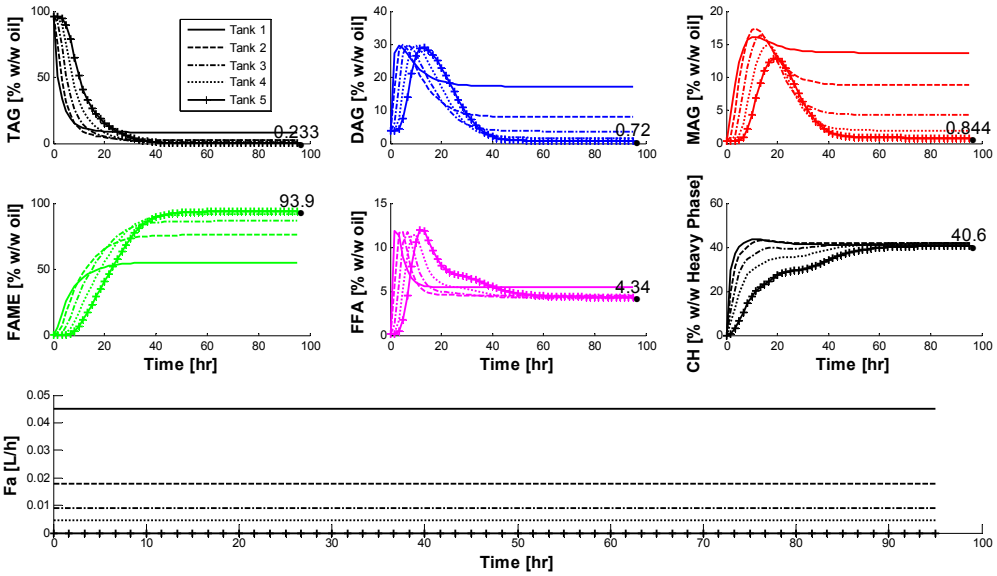


Figure 8-6 Simulation of the concentration profile of the main components in the oil for 5 CSTR's with a combined residence time of 30 hours.

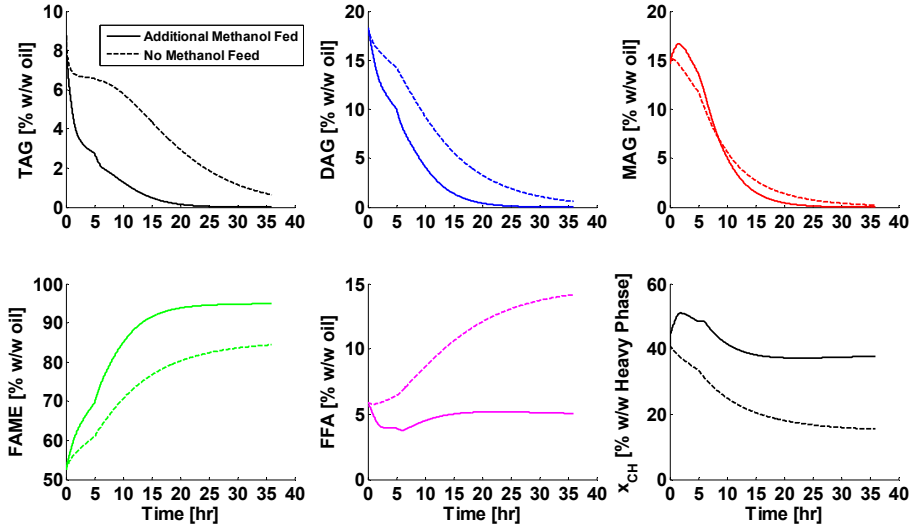


Figure 8-7 Comparison of the concentration profiles of the main components in the second tank comparing Batch (no methanol added once the tank is filled) and Fed-batch (methanol is continually added) operation.

To make a system like this work, multiple fed-batch tanks will be required. For example if the CSTR has a residence time of 5 hours and the fed-batch part of the reaction takes 25 hours then one will need a minimum of 6 additional tanks the same size as the CSTR to have smooth continuous operation. The alternative is to size the fed-batch reactor larger than the CSTR. In general terms the maximum number of tanks can be estimated by equation (8.3).

$$N_{Tank} = 1 + \left(\frac{\tau_{cstr} + t_{batch}}{\tau_{cstr}} . \frac{\tau_{cstr}}{t_f}\right) \quad \text{Eqn. 8.3}$$

Where N is the number of tanks, τ_{cstr} is the residence time in the CSTR, t_f is the filling time of the fed-batch reactor based on the flow rate out of the CSTR. If $t_f = \tau_{cstr}$ then the reactors are the same size. If t_f is greater than τ_{cstr} then less tanks are required but they need to be larger.

8.5.3. Practical implications of the Rag phase formed

A practical issue not discussed in the literature when enzymatically producing biodiesel is the formation of a Rag phase as seen in Figure 8-8. This Rag phase, can be described as a stable or semi-stable phase of emulsified reactants, which is formed during the reaction, which results in a reduction in the biodiesel yield[17]. In Figure 8-8 A) the Rag phase formed after 24 hours on the sides of the reactor during a previous fed-batch experiment was minimal. In Figure 8-8 B) the Rag phased formed during CSTR operation was quite

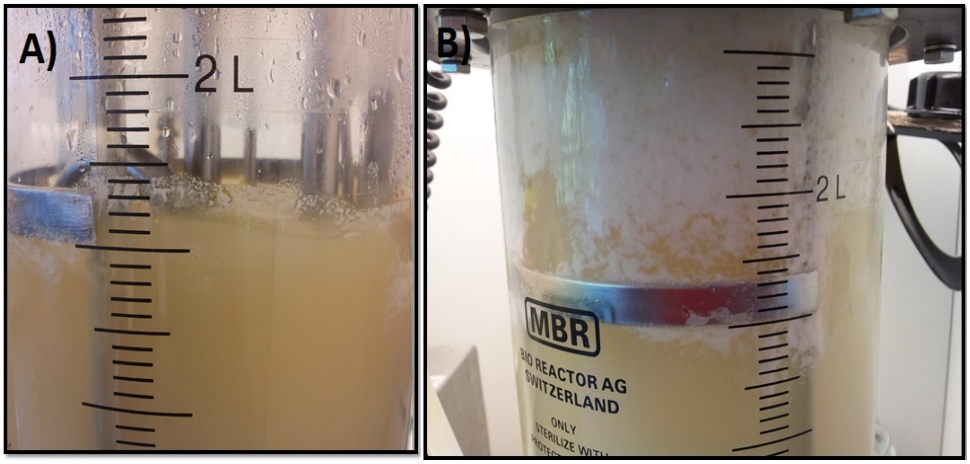


Figure 8-8 In A) (previous fed batch reaction) after 24 hours there is a minimal amount of RAG phase formed on the sides of the reactor. In B) (current CSTR operation) it can be seen that there is a substantial amount of RAG phased formed after 4 hours.

substantial only after 4 hours. It is know that monoglycerides are a good emulsifier[18]. Given the monoglycerides concentration during the CSTR operation was in the range of 10 to 20 mass % this may the reason for the increased Rag phase formed. Also, interestingly enough the plane where the Rushton turbines were located had no formation of the RAG phase on the walls of the reactor. Hence adding another impeller (which will also increase the power input) closer to the top of the reactor may have prevented the formation of the RAG phase on the walls of the reactor at the top. It should be noted the Rag layer formation reached a steady state after 4 hours. Even after 36 hours of observation the size of the layer didn't change notably compared to what is seen in Figure 8-8 B.

This Rag phases to an extent affects the economy of the process. It is know that enzyme resides in the polar phase[19]. However, some of the enzyme can be trapped in the Rag phase (personal communication, Per Munk Nielsen, Novozymes). Hence moving from fed-batch to CSTR operation will necessitate the recovery/reuse of the enzyme also from this phase.

8.6. Conclusions

The developed kinetic model is employed to carry out a study dealing with the moving from fed-batch to continuous enzymatic biodiesel production. The method of fitting fed-batch into CSTR data was superior to fitting fed-batch data alone to predict how the continuous process should be operated. Also, fitting fed-batch into CSTR data reduced the number of data points (one experiment, 240 data points) necessary to calibrate the model compared to fitting fed-batch data alone (see chapter 6, eight fed-batch experiments, 580 data points).

By manipulating the methanol feed, it is possible to have similar performance to a fed-batch operation in a CSTR. However the capital investment increases due to the number of tanks need. Compared to the fed-batch operation, the CSTR operation is much simpler and enables easier handling of cheaper, high melting point substrates. While the advantage of having multiple CSTRs means that the process can be operated continuously, taking advantage of the efficiency of a fed-batch reactor in the last half of the reaction is also an option. To achieve continuous operation in this alternative setup (CSTR into Fed-Batch) requires proper scheduling of the tanks which adds some complexity to the system, which need to be further evaluated.

8.7. References

1. Gog A, Roman M, Toşa M, Paizs C, Irimie FD. Biodiesel production using enzymatic transesterification – Current state and perspectives. *Renew Energy*. 2012;39:10–16.

2. Ganesan D, Rajendran A, Thangavelu V. An overview on the recent advances in the transesterification of vegetable oils for biodiesel production using chemical and biocatalysts. *Rev Environ Sci Bio/Technology*. 2009;8:367–394.

3. Behzadi S, Farid MM. Review: examining the use of different feedstock for the production of biodiesel. *Asia-Pacific J Chem Eng*. 2007;2:480–486.

4. Srivastava A, Prasad R. Triglycerides-based diesel fuels. *Renew Sustain Energy Rev*. 2000;4:111–133.

5. Fjerbaek L, Christensen K V, Norddahl B. A review of the current state of biodiesel production using enzymatic transesterification. *Biotechnol Bioeng*. 2009;102:1298–1315.

6. Cesarini S, Diaz P, Nielsen PM. Exploring a new, soluble lipase for FAMEs production in water-containing systems using crude soybean oil as a feedstock. *Process Biochem*. 2013;48(3):484-487.

7. Al-Zuhair S. Production of Biodiesel by Lipase-Catalyzed Transesterification of Vegetable Oils: A Kinetics Study. *Biotechnol Prog*. 2005;21:1442–1448.

8. Nordblad M, Silva VTL, Nielsen PM, Woodley JM. Identification of critical parameters in liquid enzyme-catalyzed biodiesel production. *Biotechnol Bioeng*. 2014:In Press.

9. Levenspiel O. *Chemical Reaction Engineering*. Wiley; 1999.

10. Vasić-Rački D, Findrik Z, Vrsalović Presečki A. Modelling as a tool of enzyme reaction engineering for enzyme reactor development. *Appl Microbiol Biotechnol*. 2011;91(4):845-56.

11. Brun R, Kuhni M, Siegrist H, Gujer W, Reichert P. Practical identifiability of ASM2d parameters - systematic selection and tuning of parameter subsets. *Water Res*. 2002;36:4113-4127.

12. Martinelle M, Holmquist M, Hult K. On the interfacial activation of Candida-Antarctica lipase-A and lipase-B as compared with Humicola-Lanuginosa lipase. *Biochim Biophys Acta-Lipids Lipid Metab*. 1995;1258:272-276.

13. Armenante P. Power consumption in agitated vessels provided with multiple-disk turbines. *Ind Eng Chem Res*. 1998;37(1):284 - 291.

14. Foglia TA, Jones KC. Quantitation of neutral lipid mixtures using high performance liquid chromatography with light scattering detection. *J Liq Chromatogr Relat Technol*. 1997;20:1829-1838.

15. Xu Y, Nordblad M, Woodley JM. A two-stage enzymatic ethanol-based biodiesel production in a packed bed reactor. *J Biotechnol*. 2012;162:407-414.

16. Price J, Hofmann B, Silva VTL, Nordblad M, Woodley JM, Huusom JK. Mechanistic Modelling of Biodiesel Production using a Liquid Lipase Formulation. *Biotechnol Prog*. 2014:In Press.

17. Massingill JL, Patel PN, Guntupalli M, Garret C, Ji C. High Efficiency Nondispersive Reactor for Two-Phase Reactions. *Org Process Res Dev*. 2008;12(4):771-777.

18. Hayes DG, Gulari E. 1-Monoglyceride production from lipase-catalyzed esterification of glycerol and fatty acid in reverse micelles. *Biotechnol Bioeng*. 1991;38(5):507-17.

19. Toftgaard Pedersen A, Nordblad M, Nielsen PM, Woodley JM. Batch production of FAEE-biodiesel using a liquid lipase formulation. *J Mol Catal B Enzym*. 2014;105:89-94.

Chapter 9: Addressing Plant-Model mismatch

In the final application, the imperfect model is coupled with a state estimator to correct for the process model mismatch as well as to aid in the detection of outliers in the process data.

A modified version of this chapter has been submitted for publication in the journal Biotechnology Progress as Price, J. A., Nordblad, M., Woodley, J., & Huusom, J. K. (2014). Real-Time Model Based Process Monitoring of Enzymatic Biodiesel Production.

9.1. Introduction

In the mechanistic modelling of bio-catalytic reaction, it is not unusual to have a large number of parameters and few experimental data. Usually this then means when one tries to estimate the parameters for the system, the parameters are not identifiable and model assumptions are needed to simplify the problem[1–3]. This then leads to a limited range for the predictive capabilities of the model of the bio-catalytic process.

One such bio-catalytic process of industrial relevance is the enzymatic production of biodiesel. Some of the industrial biodiesel producers, have turned to the use of a liquid lipase as a biocatalyst using (Callera™ Trans L a liquid formulation of a modified Thermomyces lanuginosus lipase), to treat a wide range of low quality/cost oil feedstock's and waste oils that have a high free fatty acid (*FFA*) content. A lipase (triacylglycerol acylhydrolase, EC 3.1. 1.3) is used, given the fact that lipases are able to esterify the *FFA* contained in waste oils to esters, as well as transesterify the acyl-glycerides in the oil; which will require additional pre-treatment steps if a conventional alkaline catalyst is used[4,5].

The formulation and validation of the mechanistic model describing the transesterification of rapeseed oil using a liquid lipase was shown in chapter 6. However, when the model is used for prediction of an optimal methanol feeding profile, the model showed poor prediction of the *FFA* concentration in the latter half of the reaction compared to the measured *FFA* value. It is not unusual when mechanistically modelling a complex system, one is not able to mathematically model all the underlying phenomena of the system. Any phenomena not modelled can lead to the model of the system potentially differing from the actual system.

The problems we have faced are not unique as seen from the various kinetic models proposed for the enzymatic transesterification of vegetable oils[6–11]. Most of the proposed models are able to capture the biodiesel concentrations over the entire course of the reaction accurately. However, the models show poor prediction of the acyl glycerides and *FFA* concentration over the entire course of the reaction. This is of great concern when using the model for predictive purposes. The reason being, that the acyl glycerides and *FFA* concentration leaving the reactor need to be within specification. If it is not within specification, it then complicates the downstream processing and it is more difficult to meet the final biodiesel fuel specification.

It is also not uncommon that the process or performance of a unit operation may change over time. For example in the biodiesel case, the enzyme loses activity over an extended period of time[5,12]. Along with the issues outlined previously is the fact that it is also usually difficult to obtain regular, noise free measurements of the process states (e.g. concentrations). The question then is how can we still use the model for reliable online prediction of the process states and monitoring of the process given the issues outlined? One possible method is the use of model based state estimation.

9.2. Model Based state estimation

In model based state estimation the states of a system are estimated using a mathematical model of the process and measurements of the system. Given measurements occur at discrete time intervals a continuous nonlinear model of the system can then be represented by the following standard discrete time equations:

$$\hat{x}_{k+1} = \hat{x}_k + \int_{t_k}^{t_{k+1}} f(\hat{x}(t), u(t))dt + w_k \text{ Eqn. } (9.1a)$$

$$\hat{y}_{k+1} = C\hat{x}_{k+1} + v_{k+1} \text{ Eqn. } (9.1b)$$

where $\hat{x}_{k+1}$ is the state prediction, at time t_{k+1} given the state of the system at the current time step $\hat{x}_k$ and the input $u(t)$ to the nonlinear model of the system and $\hat{y}_{k+1}$ is the measurement prediction at t_{k+1} where the matrix C relates the state $\hat{x}_{k+1}$ to the measurement. The model of the process and the measurements obtained are not perfect, so w_k and v_{k+1} are used to represent the process and measurement noise respectively. In this work we assume that the process and measurement noise are independent and identically distributed with a normal probability distribution:

$$p(w) = N_{iid}(0, Q)$$
$$p(v) = N_{iid}(0, R)$$

Where Q and R represents the process and the measurement noise covariance respectively.

One of the most often-used tools for stochastic estimation of states from noisy sensor measurements is the Kalman Filter[13]. The Kalman Filter is a recursive algorithm that operates on noisy input data to produce statistically optimal estimates of the underlying states of the system.[14] The algorithm uses a linear, discrete time, state space model. For a

process where we assume no disturbances, the state evolution in deviation variables can be represented as:

$$x_{k+1}^{d} = Ax_{k}^{d} + Bu_{k}^{d} + w_{k} \quad \text{Eqn. } (9.1c)$$

Where the matrix A relates the state at the previous time step to the state at the current step and matrix B relates the optional control input to the state *x.* Given the dynamics of our system is highly nonlinear, the use of a nonlinear state estimator is desirable. One possible method is the use the Extended Kalman Filter[15]. For the Extended Kalman Filter the matrices A_k and B_k is substituted for matrices A and B respectively as seen in equation (9.1d).[16]

$$A_k = \left.\frac{\partial f}{\partial x}\right|_{x_k} \; and \; B_k = \left.\frac{\partial f}{\partial u}\right|_{u_k} \quad \text{Eqn. } (9.1d)$$

For enzymatic biodiesel production, the sampling of measurements is quite infrequent. The infrequency of samples then means that the errors due to first order Taylor approximation of the nonlinear state function might have a negative influence on the accuracy of the Extended Kalman Filter.[16,17] In this work the Continuous-Discrete Extended Kalman Filter formulation is used to estimate the states of the system[15,18]. This formulation of the Extended Kalman Filter is used given that it avoids the linearization of the given nonlinear continuous-time model.

However, there have been very few applications of the Continuous-Discrete Extended Kalman Filter to biochemical reactions even though it overcomes the drawbacks outlined in regards to the Extended Kalman Filter[19–21]. The common theme from each of these applications to biochemical reactions is the ease of implementation once a nonlinear model of the system has been formulated and the ease of tuning of the state estimator to correct for process-model mismatch. Likewise, we combine our fed-batch model for enzymatic biodiesel production with knowledge of the measurement noise covariance, R and we iteratively tune the process noise covariance, Q to obtain reasonable model estimates of the measured and unmeasured states of the system. An illustration of how the process, measurements and the state estimator are coupled can be seen in Figure 9-1. The system in this case is the enzymatic fed-batch production of biodiesel. The measurement at each time step is combined with the prediction from the nonlinear model. The Kalman Filter

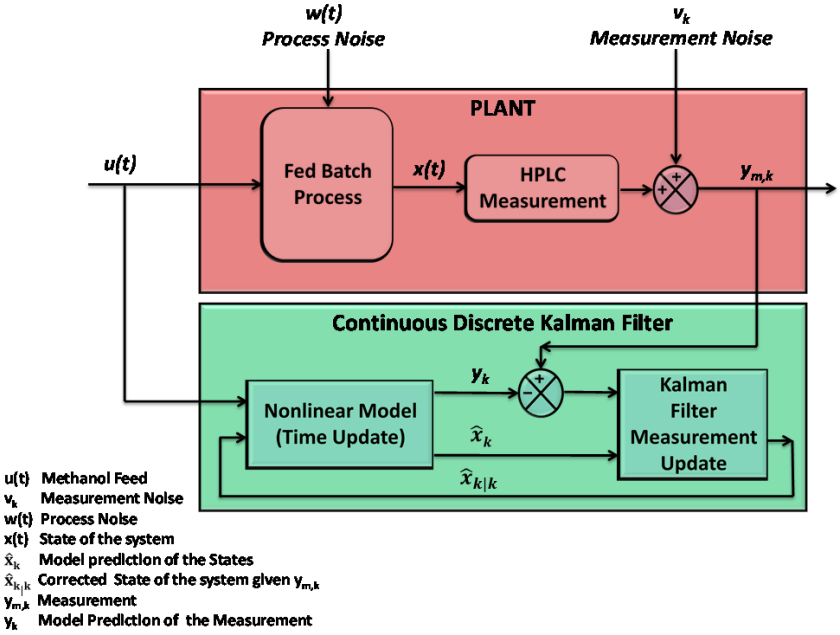


Figure 9-1 Overview of the process coupled to the state estimation

measurement update box provides an updated estimate of the states of the system by using a weighted difference between the actual measurement and the measurement prediction.

9.3. Experimental Methods and Analysis

The chemicals along with the analytical methods used can be found in our previous work.[22] The experiments used along with the measurements taken are highlighted.

9.3.1. Enzymatic Biodiesel Fed-batch Process

To test the performance of the proposed Kalman Filter, three data sets were used (see Table 9-1). These data sets cover a reasonable range of process conditions for the transesterification of rapeseed oil with methanol using the liquid lipase formulation, Callera™ Trans L. The main differences are in the amount of methanol initially dosed to initiate the reaction, the feed rate of methanol and the amount of enzyme and water used. In the first two experiment 1.5 equivalents (Eq.) of methanol was added in total. One equivalent corresponds to the stoichiometric amount of alcohol needed to convert all fatty acid residues in the oil to biodiesel (i.e. 1 mol oil : 3 mol alcohol). The third experiment is an initial rate experiment where only 0.4 Eq. of methanol is added at the start of the reaction.

Table 9-1 Conditions for the three process runs

Exp.	Mass Oil [g]	Methanol Feed Rate [Eq./h]	Initial Dose Methanol [Eq]	Water [wt.% oil]	Enzyme [wt.% oil]
1	1321	0.185 first 2hrs. 0.06 thereafter	0.2	5	0.5
2	1328	0.152 first 3hrs. 0.02 thereafter	0.525	5	0.5
3	110	0	0.4	7	0.2

9.3.2. HPLC off-line analysis

During the course of the reaction 50 µL samples were taken and prepared for off-line analysis. Analysis of the triglycerides (TAGs), diglycerides (DAGs), monoglycerides (MAGs), free fatty acids (FFAs), and fatty acid methyl esters (FAME) in the various samples were performed using an HPLC (Ultimate 3000, Dionex A/S, Hvidovre, Denmark). The composition of the reaction samples was reported on a mass percentage basis, relative to the sum of quantified mass of the five analysed components (*TAG, DAG, MAG, FFA* and *FAME*). From previous experiments the standard deviation of the measurements, σ_m for *TAG, DAG, MAG, FFA* and *FAME* were found to be 0.40, 0.75, 0.18, 0.28 and 0.26 mass % respectively.

9.4. Numerical Methods

9.4.1. Simulation environment

The model was implemented and simulated in Matlab® (The Mathworks, Natick, MA). The Continuous-Discrete Extended Kalman Filter implementation is based on the work by Jørgensen and co-workers[23,24]. The following sections give further details of the methods used.

9.4.2. Continuous-Discrete Extended Kalman Filter Algorithm

The Continuous-Discrete Extended Kalman Filter uses a nonlinear process model to compute the state and the state covariance estimates. The algorithm is comprised of two main parts, the time update equations and measurement update equations. The time update equations gives the one step ahead propagation of the a priori state ($\hat{x}_{k+1|k}$) and error covariance of the estimated states ($P_{k+1|k}$) at t_{k+1}. The measurement update

equations incorporated the new measurement (y_{k+1}) into the a priori estimate ($\hat{x}_{k+1|k}$ and $P_{k+1|k}$) to obtain an improved a posteriori estimate ($\hat{x}_{k+1|k+1}$ and $P_{k+1|k+1}$).

Selection of the values for the for the measurement noise covariance and the tuning parameter for the process noise covariance

The measurement noise covariance *R* (a square matrix the size or the number of measurements) is fairly straightforward to determine. Given measurements of the process are made, then the variance of the measurement noise can be found. *R* can then be found from equation (9.4).

$$R_{i,j} = \delta_{i,j} \text{ where } \delta_{i,j} = \begin{cases} \sigma_{m_i}^2 & i = j \\ 0 & i \neq j \end{cases} \quad \text{Eqn. (9.4)}$$

The process noise covariance *Q* (a square matrix the size or the number of states) is generally more difficult to determine given that we typically do not have the ability to directly observe the uncertainties in the process we are estimating. Hence offline tuning of *Q* is usually necessary. To simplify the procedure in the calculation of *Q* the formulation in equation (9.5) is used.

$$Q_{i,j} = \delta_{i,j} \text{ where } \delta_{i,j} = \begin{cases} q_i & i = j \\ 0 & i \neq j \end{cases} \quad \text{Eqn. } (9.5a)$$

$$q = (q_x . x_\mu)^2 \text{ Eqn. } (9.5b)$$

The tuning parameter q_x, was multiplied by x_μ which is the average value of the states over the entire reaction. This was done so as to get a reasonable scaling for q. q_x is then the parameter that was iteratively tuned where a small value of q_x was chosen (1×10^{-4}) which caused the state estimate to be the same as the pure model simulation. q_x was then was gradually increased until the measurement prediction for the state estimate and the actual measurement converged. This value of *Q* was then used as the initial value for the error covariance of the states, P.

The details of the implementation can be seen below.

Time update (one step ahead prediction)

The differential equations for the state and covariance estimate are:

$$\frac{d\hat{x}(t)}{dt} = f\big(t, \hat{x}(t), u(t)\big);\ \hat{x}(t_k) = x_{k|k} \text{ (9.6a)}$$

$$\frac{dP(t)}{dt} = A(t)P(t) + P(t)A(t)' + Q \; ; \; P(t_k) = P_{k|k} \text{ Eqn. (9.6b)}$$

where

$$A(t) = \frac{\partial f}{\partial x}\big(\hat{x}(t), u(t)\big) \text{ Eqn. (9.6c)}$$

The state and covariance estimates were then solved numerically. In this formulation of the filter, $A(t)$ represents the Jacobian matrix of the process model over the time interval t_k to t_{k+1}. Further information on the implementation can be found in the work by Kulikov and co-workers and Jørgensen and co-workers[17,24]. The integration of the state differential equations as well as the covariance differential equations were solved using a stiff variable order solver based on numerical differentiation formulas (ode15s).

Measurement update (Correction):

After the time update was performed, the measurement update equations were used to correct the state ($\hat{x}_{k+1|k}$) and covariance $P_{k+1|k}$ estimates with the measurement at y_{k+1}.

$$e_{k+1} = y_{k+1} - C\hat{x}_{k+1|k} \text{ Eqn. (9.8a)}$$

$$K_{f,k+1} = \frac{P_{k+1|k}C^T}{\left[CP_{k+1|k}C^T + R\right]} \text{ Eqn. (9.8b)}$$

$$\hat{x}_{k+1|k+1} = \hat{x}_{k+1|k} + K_{f,k+1}e_{k+1} \text{ Eqn. (9.8c)}$$

$$P_{k+1|k+1} = P_{k+1|k} - K_{f,k+1}\left[CP_{k+1|k}C^T + R\right]K_{f,k+1}^T \text{Eqn. (9.8d)}$$

where e_{k+1} reflects the discrepancy between the predicted measurement and the actual measurement. The Kalman Filter gain $K_{f,k+1}$ in equation (9.8b) is a weighting factor. As the measurement error R approaches zero, $K_{f,k+1}$ weighs the residuals more heavily (more confidence in the measurement, y_{k+1}). Whereas, if $P_{k+1|k}$ approaches zero $K_{f,k+1}$ weighs the residuals less heavily (more confidence in the model prediction, $\hat{x}_{k+1|k}$). After the Kalman Filter gain was calculated, the measurement update for the state ($\hat{x}_{k+1|k+1}$) and covariance ($P_{k+1|k+1}$) were then made. Upper and lower bounds (UB and LB respectively) for the state estimate can then be made where:

$$\hat{x}_{k+1_{UB/LB}} = \hat{x}_{k+1|k+1} \pm 3\sigma_{k+1} \text{ Eqn. (9.9}a\text{)}$$

$$\sigma_{k+1}^2 \equiv diag(P_{k+1|k+1}) \text{ Eqn. (9.9}b\text{)}$$

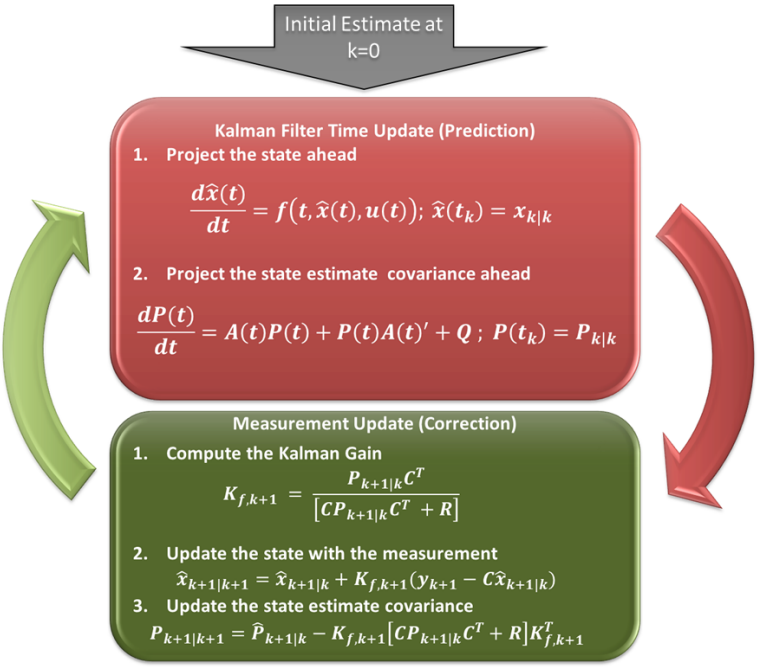


Figure 9-2 Illustration of the Continuous-Discrete Extended Kalman Filter Algorithm showing the recursive nature of the algorithm

An illustration of the Continuous-Discrete Extended Kalman Filter algorithm can be seen in Figure 9-2. Note, in the following sections, for brevity, we mean the Continuous-Discrete Extended Kalman Filter anywhere the term Kalman Filter is used.

9.5. Results and Discussion

9.5.1. Application of the filter in correcting for Process-Model mismatch on the measured states

The results comparing the predictions from the pure model simulations (integrating the nonlinear model based only on the initial conditions) compared to the state estimator predictions (use of the nonlinear model in the Continuous-Discrete Extended Kalman Filter method) are discussed in this section. The Continuous-Discrete Extended Kalman Filter does an excellent job in correcting for the process-model mismatch, over the entire time course of the reaction, for the three process runs. The results are shown in Figure 9-3 to Figure 9-5. A value of $q_x = 2 \times 10^{-2}$ was found to be able to correct for the mismatch between the model and the process data for all three process runs. It is interesting that the value of q_x found is applicable for the three different process runs with various methanol feeding

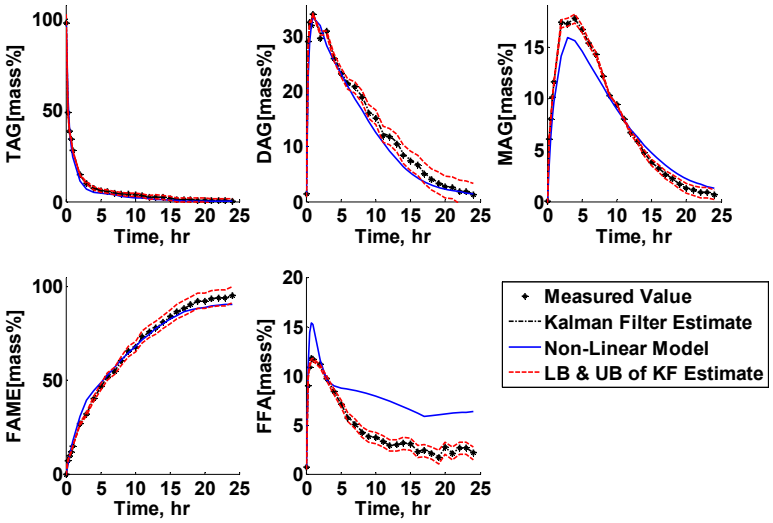


Figure 9-3 Process run 1. For this case 5 wt % water, 0.5 wt % enzyme and an initial methanol dose of 0.2 Eq methanol was used. 0.185 Eq/h of methanol is fed for 2 hours after which the feed rate was switched to 0.06 Eq/h until 1.5 Eq of methanol is added in total. The Continuous-Discrete Extended Kalman Filter estimate is for a q_x value of $2x10^{-2}$ and is compared to the measurements and the pure model simulation (nonlinear model).

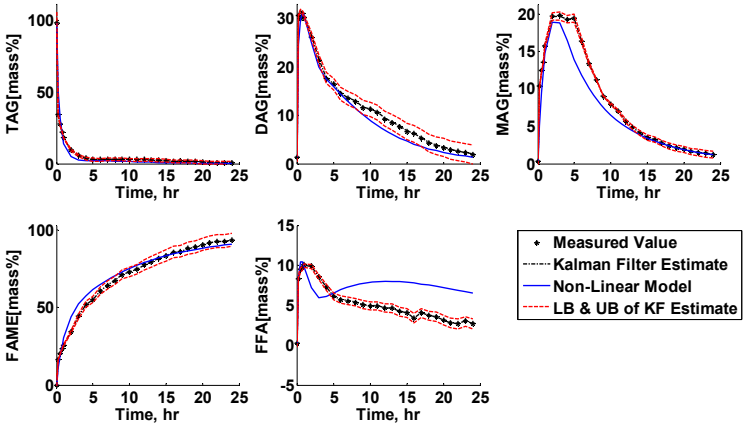


Figure 9-4 Process run 2. For this case 5 wt % water, 0.5 wt % enzyme and an initial methanol dose of 0.525 Eq methanol was used. 0.152 Eq/h of methanol is fed for 3 hours after which the feed rate was switched to 0.02 Eq/h until 1.5 Eq of methanol is added in total. The Continuous-Discrete Extended Kalman Filter estimate is for a q_x value of $2x10^{-2}$ and is compared to the measurements and the pure model simulation (nonlinear model).

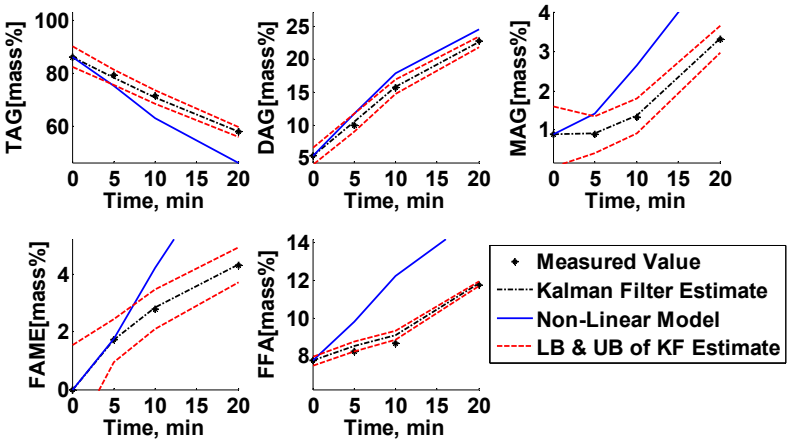


Figure 9-5 Process run 3 Initial rate experiment with an initial methanol dose of 0.4 Eq methanol , 4 wt % water and 0.3 wt % enzyme. The Continuous-Discrete Extended Kalman Filter estimate is for a q_x value of $2x10^{-2}$ and is compared to the measurements and the pure model simulation (nonlinear model).

rates, enzyme and water concentrations. The simplicity of the one tuning parameter that holds for the different operating conditions is quite powerful which enables the utilization of the model for predictive proposes given the inherent model uncertainty.

A closer look at the *FFA* plot, for process run 1 in Figure 9-3, shows the pure model prediction deviates from the measurements after five hours. The state estimator on the other hand uses the information from the Kalman Filter gain, $K_{f,k+1}$ to weight the error between the measurements received at the time to the model prediction. Upper and lower bounds are also calculated for the state estimate. These are calculated from the variance (equation (9.9b)). During the time update the variance grows (equation (9.6b)) while during the measurement update the variance shrinks (equation (9.8d)). If the variance grows more than it shrinks, then the increasing deviation in the upper and lower bounds are observed as in the case for the FAME plot in Figure 9-3. The reverse situation is seen for the MAG plot in Figure 9-3 where between 3 to 10 hrs the upper and lower bounds shrink.

The analysis performed here is done off-line. However, the results can easily be implemented for online analysis. As soon as a measurement is received, the measurement combined with the filter enables real time update of the states of the system (state estimate prediction). Compared to the pure model simulation this allows for a better prediction of the states. Also the state estimator gives the uncertainty in the prediction of the states during the time period when no measurements are taken as illustrated in Figure 9-3 to Figure 9-5.

A visual representation of the reduction in the error between the measurement and state estimate compared to the measurement and the predictions from the pure model simulations for process run 3 is Figure 9-6. It is believed that the huge deviations seen for the predictions from the pure model simulations was due to this process run being performed at higher water concentrations than what the model was calibrated to. The predictions from the pure model simulations follows the expected trends but was not very accurate. The use of the state estimator solves the accuracy issues. To get a more general perspective, the mean and standard deviation of the estimation error can be used to evaluate the statistics of the predictions from the pure model simulations vs. the state estimator predictions. This is shown in Table 9-2. There is a significant reduction in the mean and standard deviation of the estimation error for all the process runs. The mean of the estimation error for the state estimate is never zero but is significantly reduced for all the components and now the standard deviation of the estimation error for the state estimator is on the same order of magnitude as the standard deviation of the measurement error. This then gives the opportunity to use state estimator as a tool for detection of outliers in the data. Before we explore the use of the state estimator for outlier detection we investigate the effect of the state estimator on the unmeasured states.

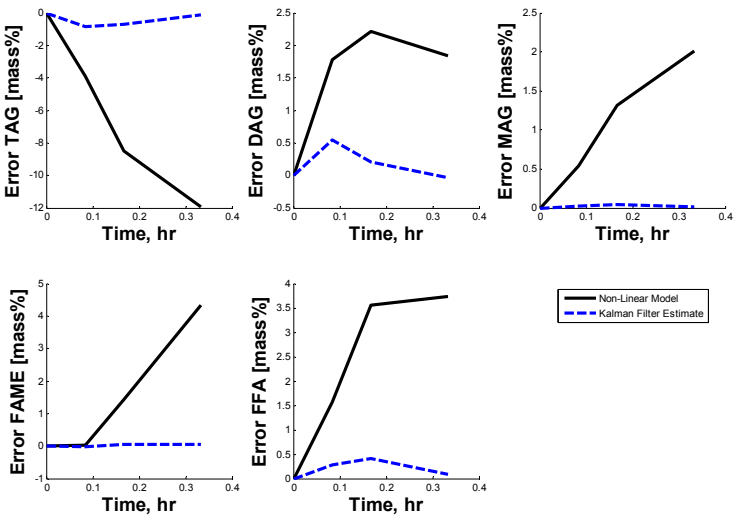


Figure 9-6 Plots showing the reduction in the Error between the Continuous-Discrete Extended Kalman Filter estimate and the measured data compared to the error between the pure model simulation (nonlinear model) and the measured data for the additional Process run 3.

Table 9-2 Comparison of the mean and standard deviation of the estimation error for the pure model simulation vs the state estimator prediction for the five measured components are reported for the three process runs. The tuning value used is $q_x = 2x10^{-2}$.

Estimation Error	Pure Model Simulation Process Run 1		State Estimate Process Run 1		Pure Model Simulation Process Run 1		State Estimate Process Run 2		Pure Model Simulation Process Run 1		State Estimate Process Run 2	
	μ_i	σ_i	μ_i	σ_i	μ_i	σ_i	μ_i	σ_i	μ_i	σ_i	μ_i	σ_i
***TAG* [Mass %]**	-0.85	2.29	-0.01	0.16	-0.58	3.73	-0.12	0.21	-6.09	5.22	-0.42	0.42
***DAG* [Mass %]**	-1.37	1.63	0.10	0.38	-1.32	1.27	0.17	0.31	1.46	0.99	0.18	0.27
***MAG* [Mass %]**	-0.47	1.35	-0.01	0.06	-1.33	1.66	-0.02	0.06	0.97	0.88	0.02	0.02
***FAME* [Mass %]**	-0.47	2.87	-0.12	0.11	1.13	3.62	-0.04	0.06	1.44	2.03	0.02	0.03
***FFA* [Mass %]**	3.16	1.38	0.04	0.19	2.10	2.11	0.01	0.07	2.22	1.77	0.20	0.19

9.5.2. Effect of the state estimator on the unmeasured states:

For the measured variables, excellent predictions are obtained from the state estimator. However, for process run 2 it was noticed that the state estimator prediction for the volume was wrong as illustrated in Figure 9-7. What is immediately apparent is that the state estimate for the change in the reactor volume (V) is grossly overestimated. During the process run the methanol addition is linear and only 0.27 L of methanol is added to the initial reactor volume of 1.6 L. The pure model prediction gives the correct time profile for the volume change as compared to the state estimate. Also a mass balance on the biodiesel measurements shows that the predictions from the pure model simulations gave smaller residuals for the methanol left in the reactor as compared to the state estimator predictions. This is due to the fact that the Kalman filter gain acts as a weighting factor (see equation (8a)-(8c)) which does not ensure that mass balance for the a posteriori state estimate ($\hat{x}_{k+1|k+1}$) closes.However, individual tuning of each diagonal element of the *Q* matrix can produce better results. For example the q_x value that relates to the volume can be set to zero. This then enables the state estimate to follow the correct evolution of the reactors' volume profile over time; given that we know the methanol feed rate to the reactor (see Figure 9-7). However, this then increases the complexity of the tuning.

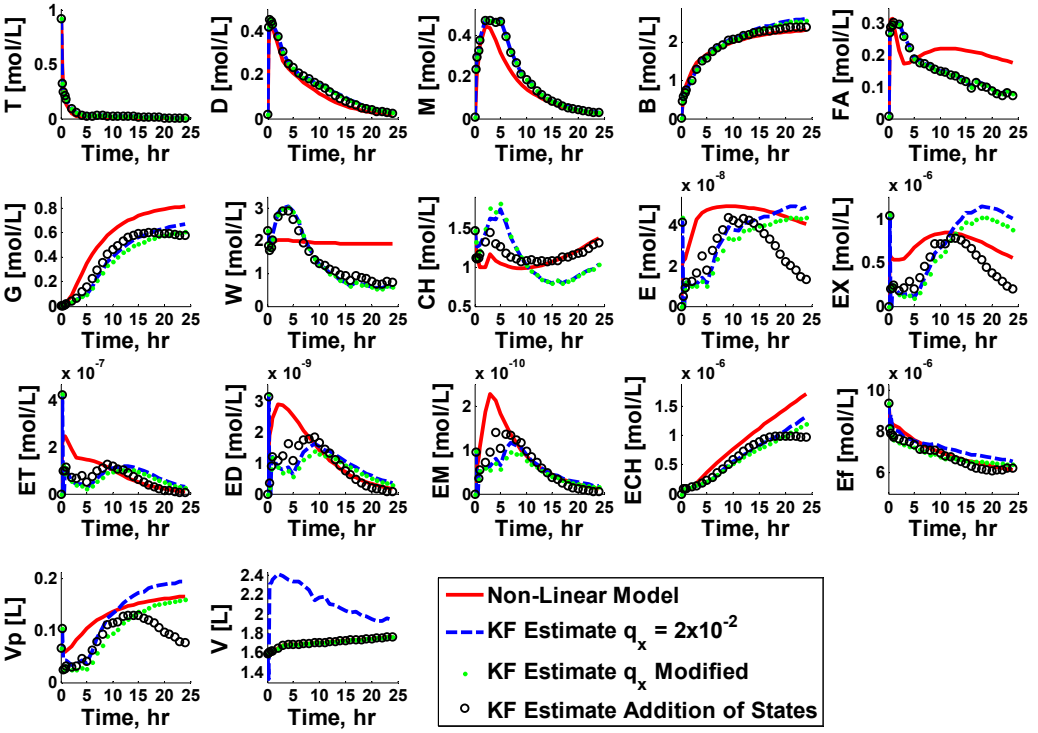


Figure 9-7 Plots of the states of the system for the nonLinear model and the Continuous-Discrete Extended Kalman Filter (KF) estimates for Process Run 2. Different modifications to the state estimate for a q_x value of $2x10^{-2}$ are investigated. Setting the q_x value that relates to the volume to zero and the addition of the states for the methanol and volume.

Another option is to use all the available information in the reconstruction of some of the unmeasured states and to add these measurements to the Kalman Filter estimation. The reactor volume as a function of time can be determined from the methanol flow rate and the methanol left in the reactor can be determined from a mass balance on the methanol fed to the reactor and the biodiesel produced. The effect of these reconstructed measurements on the performance of the state estimator can be seen in Figure 9-7 and Figure 9-8. In Figure 9-7 we now have a more realistic prediction of the volume and methanol concentration without having to tune the individual values of q_x. Comparing Figure 9-4 to Figure 9-8, it can be seen that the predictions and the variance for *TAG, DAG, MAG, FAME* and *FFA* are virtually the same. What is important is that the methanol state estimation now follows the reconstructed methanol concentration compared to the

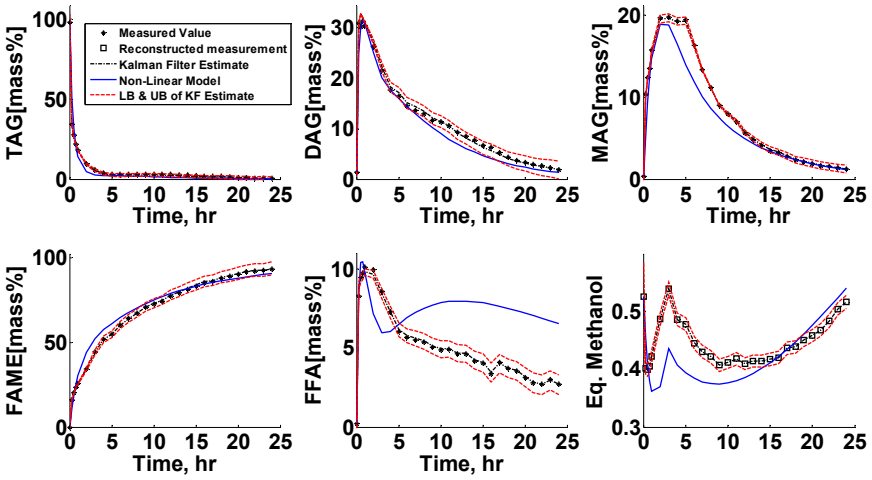


Figure 9-8 Results for the predicted measurement for Process run 2 given the addition of the states for the methanol and volume for a q_x value of $2x10^{-2}$

nonlinear model (pure model simulation) as illustrated in Figure 9-8. The advantage of using the reconstructed states is twofold. Firstly we have a better estimate of the state behaviour of the system between measurements which can be used as a monitoring tool. Secondly we can expect better online optimization of the process based on the real time state estimates. The process model can be used to optimize the feeding of methanol to the reactor[22,25]. As the reaction is proceeding, the dynamic behaviour of the system deviates from the offline model predictions as seen in Figure 4-6. Based on the estimate of the current state, as a new initial condition for the model, a new and more accurate optimization of methanol dosing for the remaining reaction time can be performed. Hence the state estimation tool enables a link between modelling and physical observation of the process, which can lead to better control and economical process operation.

Another property of the sate estimator that we wish to investigate is in the determination of outliers in our measurement data.

9.5.3. Outlier detection:

Outliers are observations that do not follow the statistical distribution of the bulk of the data.[26] The state estimate calculated (mean estimate) also has the propagated standard deviation of the mean estimate as shown by the upper and lower bounds. Values outside the upper and lower bounds give an indication of the uncertainty in the measurements and

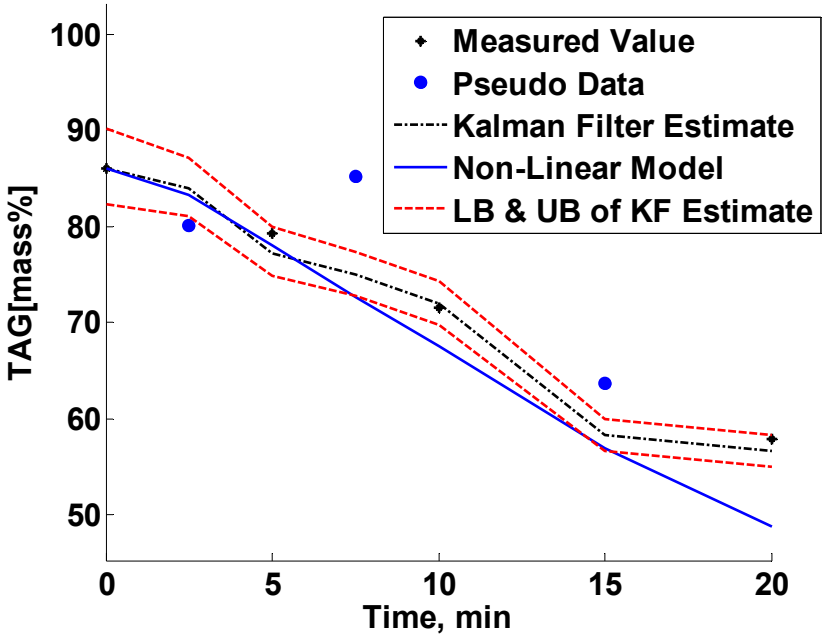


Figure 9-9 Pseudo data added to evaluate the detection of outliers

can be used as a form of measurement outlier detection. Take for example the measurement at time 10 minutes for *FFA* for process run 3 in Figure 9-5. This value falls outside the bounds and gives a visual indication of the uncertainty in the measurement.

To evaluate the state estimator as a tool for outlier detection we look at the *TAG* measurement and add pseudo data at times 2.5, 7.5 and 15 min for process run 3. The pseudo data has a standard deviation five times that of σ_m for *TAG*. The results can be seen in Figure 9-9. The outliers skew the predictions of the state estimate. This is clearly seen when the *TAG* estimate in Figure 9-5 is compared to Figure 9-9. However, the pseudo data fall outside the upper and lower bounds and is visually easy to detect.

9.6. Conclusions

There has been a general trend in mechanistic model based design, to try and describe the underlying phenomena of a process by fundamental knowledge (e.g. reaction kinetics and mass balances) of the interaction between process variables. However, it is not uncommon for slight changes to the process to render the predictive capability of the model to be inaccurate, which can lead to misleading conclusions. What we demonstrate is that

with our imperfect model, coupled to measurements of the system in the Continuous-Discrete Extended Kalman Filter framework, we can get corrected estimates of our states. The filter is relatively easy to tune given the single tuning parameter. This then lays the foundation for use of the model in a model based control framework given that it is possible to get accurate predictions of our components in the reactor, for various changes to the process inputs. This can lead to more reproducible batches and efficient utilization of methanol and the biocatalyst. Also for many processes the measurement data can be quite noisy. The state estimator can then be used to identify outliers and help filter the measurement data. The ability to correct for the process-model mismatch and identify outliers in the measurement data will prove useful in any process monitoring framework.

9.7. References

1. Jang SS, De la Hoz H, Ben-zvi A, McCaffrey WC, Gopaluni RB. Parameter estimation in models with hidden variables : An application to a biotech process. *Can J Chem Eng*. 2012;90(3):690-702.

2. Yue H, Halling P, Yu H. Model Development and Optimal Experimental Design of A Kinetically Controlled Synthesis System. In: *Proceedings of 12th IFAC Symposium on Computer Applications in Biotechnology*.; 2013:332-337.

3. Moles CG, Mendes P, Banga JR. Parameter estimation in biochemical pathways: a comparison of global optimization methods. *Genome Res*. 2003;13(11):2467-74.

4. Srivastava A, Prasad R. Triglycerides-based diesel fuels. *Renew Sustain Energy Rev*. 2000;4:111–133.

5. Fjerbaek L, Christensen K V, Norddahl B. A review of the current state of biodiesel production using enzymatic transesterification. *Biotechnol Bioeng*. 2009;102:1298–1315.

6. Lv D, Du W, Zhang G, Liu D. Mechanism study on NS81006-mediated methanolysis of triglyceride in oil/water biphasic system for biodiesel production. *Process Biochem*. 2010;45:446–450.

7. Al-Zuhair S. Production of Biodiesel by Lipase-Catalyzed Transesterification of Vegetable Oils: A Kinetics Study. *Biotechnol Prog*. 2005;21:1442–1448.

8. Cheirsilp B, H-Kittikun A, Limkatanyu S. Impact of transesterification mechanisms on the kinetic modeling of biodiesel production by immobilized lipase. *Biochem Eng J*. 2008;42:261–269.

9. Ricca E, Gabriela M, Stefano DP, Iorio G, Calabrò V, Paola M de, Curcio S. Kinetics of enzymatic trans-esterification of glycerides for biodiesel production. *Bioprocess Biosyst Eng*. 2010;33:701–710.

10. Li W, Li R, Li Q, Du W, Liu D. Acyl migration and kinetics study of 1(3)-positional specific lipase of Rhizopus oryzae-catalyzed methanolysis of triglyceride for biodiesel production. *Process Biochem*. 2010;45:1888–1893.

11. Fedosov SN, Brask J, Pedersen AK, Nordblad M, Woodley JM, Xu X. Kinetic model of biodiesel production using immobilized lipase Candida antarctica lipase B. *J Mol Catal B Enzym*. 2013;85-86:156-168.

12. Toftgaard Pedersen A, Nordblad M, Nielsen PM, Woodley JM. Batch production of FAEE-biodiesel using a liquid lipase formulation. *J Mol Catal B Enzym*. 2014;105:89-94.

13. Kalman RE. A new approach to linear filtering and prediction problems. *J basic Eng*. 1960;82(1):35-45.

14. Gelb A. *Applied Optimal Estimation*. MIT Press; 1974.

15. Jazwinski AH. *Stochastic Processes and Filtering Theory*. Academic Press; 1970:376.

16. Julier SJ, Uhlmann JK. Unscented Filtering and Nonlinear Estimation. *Proc IEEE*. 2004;92(3):401-422.

17. Kulikov GY, Kulikova M V. Accurate Numerical Implementation of the Continuous-Discrete Extended Kalman Filter. *Autom Control IEEE Trans*. 2014;59(1):273-279.

18. Zhou G, Jorgensen JB, Duwig C, Huusom JK. State Estimation in the Automotive SCR DeNOx Process. In: *Proceedings of 8th IFAC Symposium on Advanced Control of Chemical Processes*.; 2012:501-506.

19. Bogaerts P. A hybrid asymptotic-Kalman observer for bioprocesses. *Bioprocess Eng*. 1999;20(3):249-255.

20. Farza M, Hammouri H, Othman S, Busawon K. Nonlinear observers for parameter estimation in bioprocesses. *Chem Eng Sci*. 1997;52(23):4251-4267.

21. Hitzmann B, Broxtermann O, Cha Y-L, Sobieh O, Stärk E, Scheper T. The control of glucose concentration during yeast fed-batch cultivation using a fast measurement complemented by an extended Kalman filter. *Bioprocess Eng*. 2000;23(4):337-341.

22. Price J, Hofmann B, Silva VTL, Nordblad M, Woodley JM, Huusom JK. Mechanistic Modelling of Biodiesel Production using a Liquid Lipase Formulation. *Biotechnol Prog*. 2014:In Press.

23. Jørgensen JB. A Critical Discussion of the Continuous-Discrete Extended Kalman Filter. In: *European Congress of Chemical Engineering - 6*.; 2007.

24. Jorgensen JB, Thomsen PG, Madsen H, Kristensen MR. A computationally efficient and robust implementation of the continuous-discrete extended Kalman filter. In: *2007 American Control Conference, Vols 1-13*. Proceedings of the American Control Conference.; 2007:2468-2474.

25. Price J, Nordblad M, Woodley JM, Huusom JK. Fed-Batch Feeding Strategies for Enzymatic Biodiesel Production. In: *Proceedings of the 19th World Congress of the International Federation of Automatic Control*.; 2014:6204-6209.

26. Liu H, Shah S, Jiang W. On-line outlier detection and data cleaning. *Comput Chem Eng*. 2004;28(9):1635-1647.

PART IV
Discussion & Concluding Remarks

Chapter 10: Discussion

Final discussion tying the different themes of the thesis together. First an evaluation of the modeling workflow is presented followed by a discussion on the practical challenges facing enzymatic biodiesel production.

10.1. Evaluation of the modelling workflow

10.1.1. Mechanistic modelling

The aim of this thesis has been to go beyond the use of simple kinetics and use a mechanistic model based design to aid in the operation and development of an enzymatic process. In Chapter 4 the workflow for the mechanistic model development was presented. The main stages being:

- Acquiring the experimental data at relevant process conditions
- Defining the modelling objective and assumptions
- Parameter estimation
- Statistical analysis of the model to ascertain its reliability
- Use of the model

In Chapter 6 the developed kinetic model is presented along with the identifiability of the model. The structural nonidentifiability of the model is due to the fact that a change in the forward rate of reaction can be compensated for by a change in the reverse rate of reaction (highly correlated parameters). It is shown that the use of uncertainty analysis is a powerful tool to evaluate how the uncertainty in the model parameters affects the model outputs without having to make any assumptions in which parameters should be fixed while estimating others. This is actually the key in being able to use these types of models that have identifiability issues by being able to put statistical bounds on the model outputs.

While the uncertainty analysis is used to quantify the uncertainty in the model output; the use of a Continuous-Discrete Extended Kalman Filter (a state estimator) combines the mean and covariance of the error (difference between the plant and model estimate) with the model of the system to improve the model prediction (see Chapter 9). The filter is relatively easy to tune given the single tuning parameter, which makes it suited for application in a model-based control framework given the ability to get accurate predictions of the reactor components. What this means is that with the imperfect model it is now possible to get better predictions of the problematic components (eg free fatty acid).

The outcome at the end of the modelling workflow was a better understanding of the process and the ability to quickly evaluate different processing options as seen in chapter 8 where continuous operation using a soluble lipase is evaluated. These modelling tools used

in the thesis complement each other and aids in making reasonable predictions for a reaction that has multiple parallel and sequential reactions taking place. The model gives one the ability to quickly evaluate different process designs and operating strategies which drastically enhances the process development compared to running multiple experiments. There is no reason why the workflow used can't be applied to other types of similar bio-catalytic process.

However, the model development takes significant time to formulate. What needs to become commonplace is the use of model templates where once the main phenomena of a system are identified, the model building process can be much quicker. This idea is not new and is the backbone of the work by many in our research group[1–4]. The challenge is being able to judge the level of model complexity required for the modelling objective. This leads to the next point of usual contention, how useful are detailed mechanistic models given that they are usually have numerous parameters, which are usually difficult to determine with the experimental data available. Why not use a simpler model.

10.1.2. Why simple models don't "cut it"

The simplest and arguably the most well-known approaches to enzyme kinetics is the Michaelis-Menten type kinetics as depicted in equation 8.1a and 8.1b.

Enzymatic reaction

$$E + S \underset{k_r}{\overset{k_f}{\leftrightarrow}} C \xrightarrow{k_{cat}} E + P \quad (8.1a)$$

Classical Michaelis-Menten Equation

$$v(t) = \frac{k_{cat} E_o S(t)}{S(t) + K_m} \quad (8.1b)$$

$$K_m = \frac{k_{cat} + k_r}{k_f} \quad (8.1c)$$

Where E and E_o are the free and total enzyme concentration respectively, S is the substrate, C is the enzyme substrate complex, P is the concentration of the product formed, v is the initial rate and K_m is the Michaelis constant.

The Michaelis-Menten kinetic expression can be derived from the quasi-steady-state solution of a system of ordinary differential equations describing the classical system[5]. However, a system of ordinary differential equations bears an advantage over the Michaelis-Menten kinetic expression because they do not require the assumptions inherent in the quasi-steady-state approximation and the rapid equilibrium approximation. A system of ordinary differential equations avoids these assumptions so as not to bias the parameter estimation results; especially for complex kinetic schemes where the rate-determining step is not immediately apparent or for systems that involve multiple-tight binding interactions.

I will also dare say the Michaelis constant, K_m is irrelevant for characterising multi-component systems. Yes, it good for comparing enzyme formulations for single substrate reactions and can give an indication on reactor selection. For example, if K_m << S then the reaction appears to be zero order especially if the substrate is continually being fed. In which case a CSTR and Fed-batch are comparable in terms of residence/reaction time. If K_m >> S then the reaction rate appears to be first order in which case a batch reactor will have a higher productivity compared to a CSTR if based on the reaction times alone. However, K_m says nothing of how the system should be optimised. This is where a detailed model of the system shines. Also, it is common to see in the scientific literature values of K_m and the turnover number, K_{cat} of interfacial enzymes, such as lipases. However, the Michaelis-Menten model only applies to soluble enzymes and substrates present in the same phase. Expressing K_m, which has the dimension of a volume concentration, has no meaning for substrates at interfaces and should be best quantified as moles per unit area[6].

10.1.3. Extension to other bio-catalytic systems

The complication of implementing of implementing bio-catalytic processes from conventional fed-batch to continuous operation revolves around the tools available to predict how the reaction would perform in different types of reactors. In conventional catalysis, Levenspiel plots have been used for many years in reaction engineering to size and determine concentrations in various types of reactors[7]. By using batch data, plots similar to the one seen in Figure 10-1 can be made; where the inverse of the reaction rate can be plotted against the conversion. The volume of a CSTR and the volume of a plug flow reactor is then represented as the shaded areas in the Levenspiel Plots.

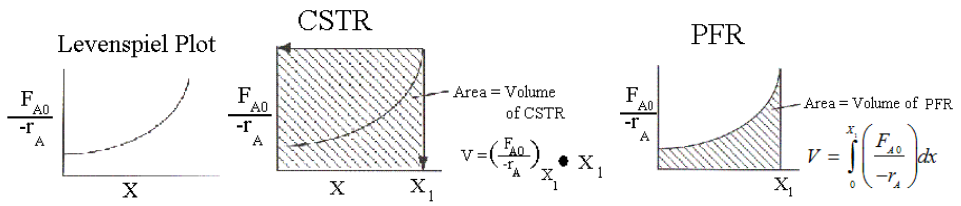


Figure 10-1 Illustration of a typical levenspiel plot where F_{Ao} if the moles of substrate A fed to the reactor r_A is the reaction rate, X is the conversion and V is the volume [7].

The use of Levenspiel plots was derived for a batch reactor, but under certain circumstances Fed-batch data can be used to approximate other reactor configurations if the rate of change in the reactor volume, dV/dt is << than the reactor volume, V. This is most relevant given many bio-processes are operated via fed-batch given the effects of substrate inhibition and it is wished to make predictions on how other reactor configurations should be operated. Also, Levenspiel plots are only for one operating condition and are not an optimization tool. This makes the mechanistic modelling approach quite attractive. For a quick guide on how a continuous process should be operated from batch data, Levenspiel plots are ideal. However, in terms of optimizing a process, mechanistic modelling will prove useful in any bio-catalytic process. With a mathematical model of the system an objective function can be formulated which can be maximised or minimized by systematically varying the input values (See chapter 7 where this is done to constrain the amount of methanol in the reactor).

10.2. Processing options

May industrial bioprocesses use fed-batch operation given that there is usually some sort of substrate inhibition. A plug flow reactor can be a possibility using a soluble lipase. However, this is not an “off the shelf” solution; given the large size needed to have the required residence time, consideration for proper mixing and inlets for the methanol dosing along the length of the reactor to minimize the deactivation of the bio-catalyst.

For the immobilised enzyme, a case can be made for using packed bed reactors such as in the production of high fructose corn syrup using immobilised glucose isomerase[8]. For the glucose isomerase system, as the activity in the reactor drops the residence time in the reactor is increased to maintain the required conversion. However, the plant productivity decreases due to the reduction in the flow-rate. However this is a very efficient way to

ensure that most of the enzyme activity is used before the spent immobilised enzyme is replaced. Compare this to the continuous enzymatic biodiesel case using a soluble lipase which is much cheaper compared to its immobilised counterpart. Since the enzyme is always being fed, to ensure that most of the enzyme activity is used, means that the enzyme needs to be recovered and recycled. This would increase the productivity requirements in terms of product produced per kg of biocatalyst. So now the trade-off is in the reduced biocatalyst cost compared to the cost of the downstream recovery of the enzyme (packed bed trade-off is the increased biocatalyst cost vs a reduction in the plant productivity; which can be mitigated by using multiple reactors with different aged catalyst). Hence for the continuous production using a soluble lipase means that the "deal breaker" is in the enzyme recovery step and this will most likely be the case for other similar types of systems. For the biodiesel case, most of the enzyme activity is in the heavy phase which can be recycled for a finite number of times before the build-up of impurities starts to affect the downstream separation process.

Nevertheless, the continuous production using a soluble lipase in a CSTR offers interesting potential for process design. For example, the use of different enzyme and/or reactions in the reactors is an interesting processing option. In this case the hydrolysis reaction can be performed in the first reactor to produce mainly free fatty acids which are then esterified in the subsequent reactors. This way enzymes that are more suited for fast hydrolysis and esterification of the oil can be used to shorten the overall reaction time. Another interesting case is in the separation of side product between reactors. Example, for the biodiesel case, glycerol can be removed between reactors to help shift the thermodynamic equilibrium and achieve better reactor productivities (Note some enzyme will also leave in this phase and would need to be recovered). Finally operating different reactors at different temperatures and methanol loadings may be a possible strategy to get the most out of the enzyme activity. This option is particularly interesting if the enzyme does not need to be recycled.

10.3. Practical challenges

In the course of carrying out the work for this thesis, there has been collaboration with various industrial collaborators such as Novozymes, Blue Sun Biodiesel and Viesel Fuel LLC . Having the privilege to work with these companies has given great insight into the actual

practical challenges faced when implementing a biocatalyst for biodiesel production. Challenges such as:

- The use of cheaper low quality feedstock's can adversely affect the final biodiesel yield
- Efficient mixing is necessary to enable efficient use of the enzyme at the oil-water interface and can be difficult to achieve when operating at large scale
- Sizing of the plant and costing of the biocatalyst so that the biocatalyst supplier and end-users business model are profitable

The general challenges in the use of low quality feedstocks, efficient mixing, sizing of the plant and costing of the bio-catalyst are not only seen in the enzymatic biodiesel process but other bio-catalytic process such as in enzymatic fat splitting or enzymatic saccharification of lignocellulosic biomass into fermentable sugars. These challenges provide excellent opportunities for research and the development of bio-catalytic processes. Our research have been leading the way in the development of metrics to aid in the costing of bio-catalytic process[9,10]. However, in terms of how to efficiently mix the reaction system and how to deal with the variability in quality of the low cost feedstock (while maintaining the final product quality) is still not a straight forward process. Both points are extremely important given that it can affect the profitability of the process; especially the latter point given that many scientific articles mention use of low cost oils can make the process more profitable but don't discuss the associated challenges of using these types of oils[11–14].

10.3.1. Feedstock variability

The use of brown grease (oil recovered from a waste water plumbing) and waste vegetable oils substantially reduce the feedstock cost. However, these types of feedstock introduce various types of impurities into the process such as emulsifiers, sulphur compounds and even some types of microorganisms. Even the use of pure vegetable oils can have variability in the free fatty acid, acylglyceride and phospholipid composition. Phospholipids act as emulsifiers and can significantly affect the downstream purification of the biodiesel if these substances are not treated. High concentrations of these emulsifiers form quite stable emulsions that can easily triple the separation time (from 3 hours to over 10 hours) of the fatty acid methyl esters from the heavy phase (water, methanol and glycerol). As it stands, the model can account for variability in the acylglycerides and fatty

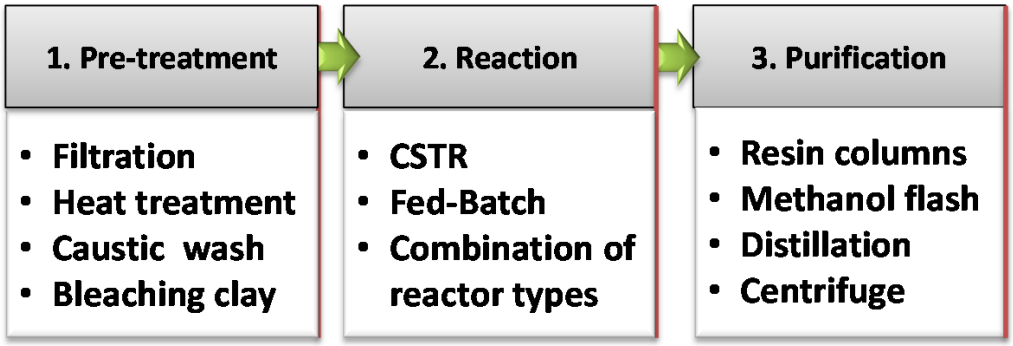


Figure 10-2 Illustration of various Process plant sections and different unit operations that could be used when using different types of feedstock's for biodiesel production

acid concentration. However, the effect of these extra components in the oil on the process still needs further investigation.

Given that the enzymatic biodiesel reaction is carried out at 35 ^{o}C, blending of the different types of feedstock is necessary to ensure that the oil is a homogeneous liquid. The use of a CSTR as the initial reactor offers the advantage that the incoming feed which may not be soluble at 35 ^{o}C can be preheated until it is liquid before it enters the first reactor. The first reactor would have fatty acid methyl esters already formed which acts as a solvent helping to solubilise the incoming feed.

In chapter 2 an overview of the processing steps for biodiesel production was given. However, this was for vegetable oils, where the pre-treatment mainly focused on degumming and reduction of free fatty acids. Taking inspiration for the oil and natural gas industry the process plant can be divided into the pre-treatment, reaction and purification section as illustrated in Figure 10-2. As can also be seen in the figure there are different unit operations that could be used to achieve a particular specification entering into each section. Formulating the plant like this gives the plant operators much more flexibility in the feedstock selection, while still maintaining the final product specification.

10.3.2. Mixing

Various correlations for the mixing and interfacial area for enzymatic biodiesel production has ben proposed[15,16]. These correlations which work well in the lab have yet to be extended for use in large scale reactors. Also the ability to predict how impurities in the

oil can affect the quality of the oil water interface is not a trivial task. Hence, screening and testing of new batches of oil are essential to determine if the pre-treatment step can handle the incoming impurities; so that the reaction and purification steps are able to perform optimally.

It is also important that there is proper mixing with no dead spots in the reactor and that reasonable recirculation times in the reactor can be achieved. The use of hydrodynamic mixing (use of pumps, ejectors and inductors for mixing) appears to be the preferred technology for mixing at scale for enzymatic biodiesel production. The use of computational fluid dynamics can be a quite useful tool to aid in determination of the best placement of the inductors and ejectors for the system given that most of the mixing design is currently done by trial and error.

The use of ultrasonic mixing is also another technology under investigation which can aid in increasing the available oil water interface. It is known that ultrasonic mixing can increase the reaction rate by increasing the interfacial area[17]. However, the power inputs to these systems are not reasonable (on the order of 100 W/L). While hydrodynamic mixing will give good mixing (residence time distribution close to that of an ideal system). It may not form as good as an emulsion as the ultrasonic mixer. In which case, a combination of the different types of mixers may be a superior option. What still needs to be determined is how does the droplet sizes distribution and residence time distribution for a particular power input affect the corresponding yield for a given enzyme loading.

10.4. References

1. Heitzig M, Sin G, Sales-Cruz M, Glarborg P, Gani R. Computer-Aided Modeling Framework for Efficient Model Development, Analysis, and Identification: Combustion and Reactor Modeling. *Ind Eng Chem Res*. 2011;50:5253–5265.

2. Mansouri SS, Ismail MI, Babi DK, Simasatitkul L, Huusom JK, Gani R. Systematic Sustainable Process Design and Analysis of Biodiesel Processes. *Processes*. 2013;1(2).

3. Al-Haque N, Santacoloma PA, Neto W, Tufvesson P, Gani R, Woodley JM. A robust methodology for kinetic model parameter estimation for biocatalytic reactions. *Biotechnol Prog*. 2012;28:1186–1196.

4. Singh R, Gernaey K V., Gani R. Model-based computer-aided framework for design of process monitoring and analysis systems. *Comput Chem Eng*. 2009;33(1):22-42.

5. Chen WW, Niepel M, Sorger PK. Classic and contemporary approaches to modeling biochemical reactions. *Genes Dev*. 2010;24(17):1861-75.

6. Reis P, Holmberg K, Watzke H, Leser ME, Miller R. Lipases at interfaces: a review. *Adv Colloid Interface Sci*. 2009;147-148:237-50.

7. Levenspiel O. *Chemical Reaction Engineering*. Wiley; 1999.

8. Schnyder BB. Continuous lsomerization of Glucose to Fructose on a Commercial Basis. *Starch*. 1974;26(12):409-412.

9. Tufvesson P, Lima-Ramos J, Nordblad M, Woodley JM. Guidelines and Cost Analysis for Catalyst Production in Biocatalytic Processes. *Org Process Res Dev*. 2011;15:266-274.

10. Tufvesson P, Lima-Ramos J, Al Haque N, Gernaey K V, Woodley JM. Advances in the Process Development of Biocatalytic Processes. *Org Process Res Dev*. 2013;17:1233-1238.

11. Samukawa T, Kaieda M, Matsumoto T, Ban K, Kondo A, Shimada Y, Noda H, Fukuda H. Pretreatment of immobilized Candida antarctica lipase for biodiesel fuel production from plant oil. *J Biosci Bioeng*. 2000;90:180–183.

12. Al-Zuhair S, Dowaidar A, Kamal H. Dynamic modeling of biodiesel production from simulated waste cooking oil using immobilized lipase. *Biochem Eng J*. 2009;44:256–262.

13. Al-zuhair S, Almenhali A, Hamad I, Alshehhi M, Alsuwaidi N, Mohamed S. Enzymatic production of biodiesel from used/waste vegetable oils: Design of a pilot plant: Renewable Energy: Generation & Application. *Renew Energy*. 2011;36(10):2605–2614.

14. Cesarini S, Diaz P, Nielsen PM. Exploring a new, soluble lipase for FAMEs production in water-containing systems using crude soybean oil as a feedstock. *Process Biochem*. 2013;48(3):484-487.

15. Al-Zuhair S, Ramachandran KB, Hasan M. Investigation of the specific interfacial area of a palm oil-water system. *J Chem Technol Biotechnol*. 2004;79:706-710.

16. Al-Zuhair S, Hasan M, Ramachandran KB. Kinetics of the enzymatic hydrolysis of palm oil by lipase. *Process Biochem*. 2003;38:1155-1163.

17. Huang J, Liu Y, Song Z, Jin Q, Liu Y, Wang X. Kinetic study on the effect of ultrasound on lipase-catalyzed hydrolysis of soy oil: Study of the interfacial area and the initial rates. *Ultrason Sonochem*. 2010;17(3):521-5.

Chapter 11: Concluding Remarks and Future Perspectives

Concluding remarks and future perspectives for model based design of enzymatic processes

11.1. Conclusion

In this thesis the use of model based process design to aid in the operation and development of enzymatic biodiesel production was explored. A short review of the biodiesel process and the liquid lipase used to catalyse the process was elaborated on. The workflow for the development of the model for the system was explained and the model then applied to evaluate how the process should be operated in fed-batch and continuous operation.

Through the modeling workflow process it is shown how such a mechanistic model can be developed for a bio-catalytic system going beyond the use of simple Michaelis-Menten kinetics. While the kinetic model is complex, it is shown how the model can still be used given the uncertainty in the model parameters to guide process development. For the uncertainty analysis, the Monte Carlo procedure was used to statistically quantify the variability in the model outputs due to uncertainties in the kinetic parameter estimates. The uncertainty analysis is a powerful tool that aids in the decision making process by being able to ascertain how reliable the model is to uncertainties in the model parameters.

The model was first used for fed-batch operation to determine an optimal methanol feeding profile that constrained the amount of methanol in the reactor, which was also experimentally validated. The main disadvantage of fed-batch operation is the downtime between batches, which reduces the reactor productivity. To address this, the model was then used to guide process development of a continuous enzymatic biodiesel process to determine reactor residence times for a desired conversion. The challenge in applying a detailed mechanistic model is that given the large number of parameters and the often few experimental data points, the parameters found are not identifiable. The model is then only applicable within the operating range for which the model was validated. The fitting of the model to fed-batch and continuous stirred tank reactor (CSTR) data, caused a reduction in the correlation between the parameters and the model described the validation dataset for the five measured components (triglycerides, diglycerides, monoglycerides, free fatty acid and fatty acid methyl esters (biodiesel)) much better than using fed-batch data alone. For the simulation case the model predicts that 5 reactors will be needed with a combined residence time of 30 hours to give comparable performance to a fed-batch operation which lasts 24 hrs assuming a 6 hour emptying and filling time at the industrial scale for the fed-

batch reaction. While the advantage of having multiple CSTRs means that the process can be operated continuously, taking advantage of the efficiency of a fed-batch reactor in the last half of the reaction is also an option and it has been shown how such a system can be operated in Chapter 8. In general, the mechanistic model allows us to evaluate the technical feasible of a continuous process using a soluble lipase for enzymatic biodiesel production. The use of a soluble lipase formulation significantly decreases the cost of the biocatalyst and improves the economics of the enzymatic biodiesel process. However, what will make the process profitable depends on the downstream process. More specifically, the ability to efficiently recover the enzyme and the ability to achieve the required biodiesel specification.

While the model seems to be "fit for purpose", it does not necessarily represent reality, as reality is much more complex. Also, the performance of a unit operation may change over time (e.g. activity of the enzyme). What then happens is that there is a mismatch between the process data and model of the system. It is shown in this work that by using a Continuous-Discrete Extended Kalman Filter (a state estimator) the process-model mismatch can be corrected. It was possible to use one tuning parameter, $q_x = 2 \times 10^{-2}$ (q_x represents the uncertainty in the process model) to reduce the overall mean and standard deviation of the error between the model and the process data for all of the five measured components (triglycerides, diglycerides, monoglycerides, fatty acid methyl esters and free fatty acid) over the entire course of the reaction. It is also shown that the state estimator can be used as a tool for detection of outliers in the measurement data. For the enzymatic biodiesel process, given the infrequent and sometimes uncertain measurements obtained, the use of the Continuous-Discrete Extended Kalman Filter is seen to be a viable tool for real time process monitoring.

What this thesis adds to the current state of the art for enzymatic biodiesel production is the use of the mechanistic model in process design. While the whole process of mechanistic modeling of a bio-catalytic process can take up much of the allocated time for a project; it causes one to dig deep into the underlying phenomena of the system, which in itself was a valuable exercise. The workflow was successfully applied to the lipase-catalyzed biodiesel production to predict how the process should be operated and for process design. It is envisaged that the methods and tools used in the workflow for the biodiesel case study can

be applied to other bio-catalytic process to assist in understanding of the process and process development.

11.2. Open Challenges and Future perspectives

Due to time limitations, there were aspects in the thesis that have not been fully concluded. One limitation I would greatly like to address if given more time is to investigate the types of experiments needed to achieve "good" parameter estimates in complex mechanistic models. By using the differences in the reactor mass balance (Fed-batch into CSTR operation shown in chapter 8) it was shown that the correlation between the highly correlated parameters were reduced. However, this need to be evaluated for other systems (This would entail more experiments using a different system not only bio-catalytic). If the same results are obtained, then further analysis of the underling theory of why this is so, would be needed to help in the development of the method. Also, being able to experimentally validate the CSTR simulations in chapter 8 would give further credence to use of mechanistic modelling in process design and operation.

Likewise, during the research project, there were many interesting paths that the research may have taken. Now that it has been concluded, there are definitely areas that could be extended upon. These areas will be divided in two areas, process research related to bio-catalytic processes and research related to the modelling of bio-catalytic process.

Process research

As mentioned in the discussion chapter, ultrasonic mixing is a promising process technology for formation of an emulsion. However, the evaluation of the effect of ultrasonic mixing on the activity of the lipase at industrially relevant power inputs still need to be determined. Likewise, it also needs to be determined what is the most efficient way to combine ultrasonic mixers with conventional mixers to ensure that the residence time distribution is close to that of an ideal system.

Turning focus to the enzyme, research on the immobilization of enzymes on magnetic nanoparticles seems like a promising technology. This should aid in the efficient recovery of the enzyme. The challenge is to create an immobilization procedure that doesn't substantially increasing the cost of the bio-catalyst.

In terms of monitoring and control, in Chapter 9, the use of state estimation theory proved to be powerful tool in correcting for the mismatch between the process data and the model. Evaluation of combining the state estimator with measurements such as viscosity to predict the concentrations of the components in the reactor would be a relatively cheap way to have real time monitoring of the process. Having real time predictions of the states of the system is even more important when moving to continuous operation so that timely adjusts can be made to the system, to ensure that the steady state of the system is maintained.

Modelling

In the modelling of lipase-catalyzed reactions it is not always clear which concentrations should be used in the kinetic equation. Should it be the concentration in the aqueous phase, in the organic phase or at the interface? The use of thermodynamic activities may be one way to circumvent this issue, given that at equilibrium, the thermodynamic activity of a component in the system is equal in all phases. However, the main drawback is in the increased complexity in the modelling and the availability of robust thermodynamic models to predict the activities of the components in the reaction.

In the model development, enzyme deactivation was neglected given that it was assumed that for the operating conditions used, that the enzyme deactivation was negligible. However, to evaluate the long term use and stability of the enzyme the addition of enzyme deactivation kinetics to the developed mechanistic model is necessary. Building on the need to add enzyme deactivation kinetics to the mechanistic model is the ability to design experiments to uniquely identify the inhibition constants from the deactivation constants.

The enzymatic biodiesel case study proved to be an interesting case study in the evaluation of reactor operation using a soluble lipase. The success of the enzymatic biodiesel process is foreseen to grow over the coming years. This will help to build confidence in the industrial sector for the application of biocatalysts in industry, for cases where biocatalyst are advantageous.

CAPEC-PROCESS
Computer Aided Process Engineering/
Process Engineering and Technology center

Department of Chemical and Biochemical Engineering
Technical University of Denmark
Søltofts Plads, Building 229
DK-2800 Kgs. Lyngby
Denmark

Phone: +45 4525 2800
Fax: +45 4525 2906
Web: www.capec-process.kt.dtu.dk

ISBN : 978-87-93054-58-5